Paris 1 - Panthéon Sorbonne University

Autonomous University of Madrid

the 5th of June 2013

PhD in Economics

Emmanuelle LAVAINE

Social inequalities linked to the effect of pollution on health

Thesis director

Mireille Chiroleu Assouline	-	Professor at Paris 1 Panthéon-Sorbonne University
Carmen Arguedas Tomás	-	Professor at Autonomous University of Madrid

Contents

1. This chapter is a joint work with Matthew Neidell from Columbia University

Introduction

Many pollutants are declining throughout the industrialized world. However, exposure to air pollution, even at the levels commonly achieved nowadays in European countries, still leads to adverse health effects. In this context, there has been increasing global concern over the public health impacts attributed to environmental pollution.

We propose to investigate the causal effect of air pollution on infants' health, respiratory outcomes in France using several natural experiments and a unique dataset combining data on environmental quality, health and property prices. Recently, the role of clean air policies have been increasing along with the rise of public health concerns in Europe. Given this increasing needs worldwide, further studies of air pollution policies are useful in order to better inform this important public policy process.

The first objective of the thesis is to explore empirically the relations between socio-economic status, environmental exposures and health outcomes. We also went further in the analysis of social inequalities linked to environmental pollution by sheding light on their macroeconomic consequences. I study differences in exposition and underline their consequences on morbidity by developing a small area empirical approach. First, I empirically measure the impact of pollution on health and productivity and look at how pollution can contribute to health inequalities. To do so, I estimate the relationship in France between nitrogen dioxyde (NO2), environmental disparities and non incidental mortality rates. This study is part of new research on environmental justice, and provides an overview of the distribution of environmental risks. Second, we estimate the health externalities from oil production by exploiting the oil refinery strikes in France in October 2010. The strikes provide a natural experiment that enables us to overcome the typical omitted variable bias that arises from Tiebout sorting (Banzhaf and Walsh, 2008) (Greenstone 2003). Amid nationwide protests over pensions reform and broader concerns about oil industry practices in France, striking workers blocked refineries, which resulted in a complete brought cessation of operations to a halt at several major refineries for nearly a month. As we demonstrate, this lead to a sharp reduction in SO2 that which quickly dissipated rose again once the strike was resolved and production resumed. We exploit this temporal event to identify the infant health externalities at birth from oil production, comparing outcomes in areas close to the refineries before and after the strike vs. during the strike, using areas far from the refineries as a control group. Finally, I have been developing a third research project focusing on the link between respiratory outcomes, housing prices and pollution using the hedonic price method. I try to draw inferences about individuals' valuations of risk by combining estimates of the effect of air pollution on both property values and hospital respiratory admissions for respiratory causes in France. The analysis focuses on Dunkerque, a french census track in the Nord-Pas-de-Calais region in France where residents have recently experienced a refinery closure. Housing prices are compared before and after the closure with the nearby census tracks within 50 kilometers, acting as a control group.

Chapter 1

Atmospheric Pollution, environmental disparities and Mortality Rate: An Econometric Analysis

This paper presents the first study of environmental inequality related to health in France on the national scale. Through an econometric analysis based on panel data from 2000 to 2004 at departmental level, I investigate the total mortality rate in relation to socioeconomic status and air pollution. The concentration level of NO_2 , O_3 and PM_{10} are estimated by spatial interpolation from local observations by a network of monitoring stations. I find a positive and significant relationship between NO_2 and the mortality rate, at mean levels below the current standard, with a greater relative risk for women. Moreover I observe disparities in health through income among French departments. These results not only confirm the existence of a relationship between current air pollution levels and mortality but also raise questions about environmental policy implications in France.

1.1. Introduction

Many pollutants are declining throughout the industrialized world. However, exposure to air pollution, even at the levels commonly reached nowadays in European countries, still leads to adverse health effects. In this context, there has been increasing global concern over the public health impacts attributed to environmental pollution.

Multilevel modelling has been previously used to assess the negative correlation between pollution exposure and socioeconomics status, such as unemployment, education, and the working class in Canada (Premji *et al.*, 2007), ethnic group, and income in England (McLeod *et al.*, 2000) and in the US (Grineski *et al.*, 2007), (Morello Frosch *et al.*, 2002) where the concept of environmental justice has been the object of increasing attention. Viel et al. emphasize that towns with high proportions of immigrants tend to host more hazardous sites even when controlled for population size, income, degree of industrialization of the town, and region (Viel *et al.*, 2011). In Germany, Schikowski et al. show the existence of social differences in respiratory health among the female population (Schikowski *et al.*, 2008) and Bolte et al. acknowledge social inequality in perceived environmental exposure in relation to housing conditions (Bolte *et al.*, 2010). Pearce et al. for New Zealand point out that industrial pollution is greater in wealthy places, whereas overall pollution affects poorer zones more (Pearce *et al.*, 2010) .

Moreover, multiple models also estimate the relationship between health and pollution, showing the impact of outdoor air pollution on the mortality rate in Austria, France and Switzerland (Kunzli *et al.*, 2000), in England (Janke *et al.*, 2009), on the allergic sensitization on primary schoolchildren in France (Maesano *et al.*, 2007), on asthma (Wilhelm et al. 2009), or on cancer risks among schoolchildren in the US (Chay *et al.*, 2003), (Morello Frosch *et al.*, 2002). Finally, Finkelstein et al. point out that mean pollutant levels tend to be higher in lower income neighbourhoods in Ontario and both income and pollutant levels are associated with mortality differences (Finkelstein *et al.*, 2003).

In addition, the literature is far to be silent about the relationship between health and socioeconomic status (SES). A number of SES measures have been proposed, including income, wealth, education, labor force status, and race/ethnicity. For instance, some recent papers provide evidence of a positive association between income and health (Subramanian & Kawachi, 2000), (Gunasekara *et al.*, 2011). Apart from income, education has also been considered a crucial component of SES affecting health (Grossman, 2000). In France, Cambois & Jusot (2010) study the link between lifelong adverse experiences, health and SES. Lifelong adverse experiences is related to poor self-perceived health, diseases and activity limitations even controlling for SES. Results from Stringhini *et al.* (2012) suggest that the social patterning of unhealthy behaviors differs between countries. They stress health behaviors are likely to only be major contributors of socioeconomic differences in health. Among others, Lindahl (2005) focus on mortality rates and find a positive causal relationship between income and health measure.

Moreover, I observe a growing epidemiologic literature about the effects of air pollution on health by gender. The most recent gender analysis from Clougherty shows that most studies for adults report stronger effects among women, particularly when using residential exposure assessment (Clougherty, 2010). The smaller size of the trachea has been argued to be a reason which makes women more sensitive to particulates inthe air (Marr, 2010). However, it remains unclear whether the observed difference is a result of gender-linked biological differences or gender differences in activity patterns.

The analysis offers several contributions to the existing literature. Most international empirical economic studies estimate either the relationship between health and pollution or the correlation between pollution exposure and socioeconomic status. I aim to gather both literatures to assess the impact of air pollution on health according to social status. Few European studies investigated the effect modification of socio-economic factors on the association between air pollution and health and much is yet to be understood (Deguen & Zmirou, 2012). European policy-makers have in fact only recently acknowledged the notions of environmental justice and environmental inequalities, which have been part of the US policy arsenal for almost two decades (Laurent, 2011). To my knowledge, environmental factors affecting health, such as exposure to atmospheric air pollution have not been yet studied in France on a national scale in the context of social inequalities. Laurent et al. emphasize the importance of continuing to investigate this topic due to the tendency for greater effects to be observed among the more deprived (Laurent *et al.*, 2007). Whereas the french literature only looks at high level of pollution, I am studying ambient air pollution; the dataset presents low level of pollution concentration, below the actual threshold fixed by the public authorities at which health can be harmed (Pascal *et al.*, 2009). Instead of looking at one geographical area, I examine recent relationships between pollution and health for the entire country using a panel dataset. I also account for unobserved confounders using fixed effects clustering at the regional level not to suffer from potential omitted variable bias. Most of the studies on this topic use times series or a cross sectional cohort (Janke *et al.*, 2009). Times series exploit short-term variation to identify pollutant effects which eliminates the effects of lifestyle factors such as smoking, exercise and diet, because these factors do not change on the short run. The cohort studies may also suffer from omitted variables bias, as the cities or zip codes which are compared, may differ from each other in important ways other than just their levels of pollution. Some recent studies use a exogeneous event to cope with omitted variables biais (Chay & Greenstone, 2003) (Moretti & Neidell, 2011), (Currie & Walker, 2011). For instance, Chay and Greenstone use a sudden recession as an instrument to identify the effect of a medium-term reduction of pollution on infant mortality (Chay & Greenstone, 2003). Finally, I use a model which takes into account spatial autocorrelation.

This paper investigates the relationship between ambient air pollutant concentrations, social class, and population mortality on the departmental scale in France [1]. It is part of new research on environmental justice, and provides an overview of the distribution of environmental risks. To identify the social distribution of air pollution, the study compared the social characteristics (income, unemployment) and the concentration of air pollution among French local authorities for different level of poverty. In this context, we first may wonder whether poor areas are also the ones with low socioeconomic level. Poor people may be more likely to live where pollution may be higher, next to industrial area (Mohai *et al.*, 2009). Due to budget constraint, the unemployed people are also less likely to move from one area to another to avoid pollution. Secondly, we also ask oneself if a change in pollution

1. Department corresponds to a local authority below the regional level.

benefits, in term of health, even more to high socioeconomic than low levels areas. When it comes to poor local authorities, is the health effect of an increase in air pollution twice over? The main purpose is to figure out if inequalities tend to mount up within French local authorities.

I find a positive and significant relationship between NO_2 and the mortality rate, at mean levels below the current standard, with a greater relative risk for women. I show higher is the income level for French department, lower is the level of mortality rate. However, health disparities appear to be more related to socioeconomic factors than differences in sensitivity to pollution.

1.2. Medical perspective

The L.A.U.R.E (Law on Air and Rational Use of Energy) and the different European directives give priority to monitoring common air pollutants with a direct effect on health, such as Nitrogen Dioxide (NO_2), Nitrogen Oxides (NO), Ozone (O_3) and Particles (PM_{10}). In consequence, I consider these pollutants in this paper. The contamination of the atmosphere by pollutants at the local and regional level is the result of three processes: emission, transmission, and air pollution concentration. Pollutants are first released at the source with gases and particles which are put into the air. The pollutants emitted are then dispersed, or sometimes they can be chemically transformed in the atmosphere, creating new, secondary pollutants. Having combined with air and become diluted, they create a concentration of toxic levels of chemicals in the air, and these atmospheric pollutants are finally inhaled by humans, animals and plants.

First, Particulate Matter (PM) is made up of a number of components, including acids, organic chemicals, metals, and soil or dust particles. The size of particles is directly linked to their effect on health: PM_{10} (aerodynamic diameter less than 10 μm); $PM_{2.5}$ (aerodynamic diameter less than 2,5 μm) are the particles that generally pass through the throat and nose and enter the lungs. The $PM_{2.5}$ particles are the most dangerous. The effects of PM on health occur at levels of exposure currently being experienced by most urban and rural populations in both developed and developing countries. Chronic exposure to particles contributes to the risk of developing cardiovascular and respiratory diseases, as well as of lung cancer (WHO). Once inhaled, these particles can affect the heart and lungs and cause serious health effects. Not only have many European projects found a link between particles and mortality or morbidity (Peng *et al.*, 2004) (Touloumi *et al.*, 1997), but so have recent epidemiologic studies (Schikowski *et al.*, 2008), (Janke *et al.*, 2009), (Maesano *et al.*, 2007).

Nitrogen Oxides (NOx) is the main indicator of transportation and stationary combustion sources, such as electric utility and industrial boilers contamination[2].

2. The spatial distribution of NO2 is generally not homogeneous within individual metropolitan areas. The primary reason for the observed heterogeneity in concentrations across an urban area is the substantially higher concentrations of NO2 near sources, such as roads [Electric Power Research Institute 2009].

NOx forms when fuels are burned at high temperatures and includes various Nitrogen compounds such as Nitrogen Dioxide (NO_2) and Nitric Oxide (NO). Nitrogen Dioxide (NO_2) and Nitric Oxide (NO) play crucial role in the atmospheric reactions by creating harmful particulate matter, ground-level Ozone, acid rain, and eutrophication of coastal waters. NO_2 is produced by chemical transformation with NO and Ozone ($NO + O_3 = NO_2 + O_2$). Not only particle filters but also the rise of Ozone in the atmosphere increase NO_2 emissions (AFSSET). As a consequence, NOx is a powerful oxidizing gas, linked with a number of adverse effects on the respiratory system (Agency, 2011).

Ozone (O_3) is an example of a secondary pollutant as it is formed when Hydrocarbons (HC) and Nitrogen Oxides (NOx) combine in the presence of sunlight. And excessive Ozone in the air can have a marked effect on human health. It can cause breathing problems, trigger asthma, reduce lung function and cause lung diseases (OMS). Breathing ozone can trigger a variety of health problems including chest pain, coughing, throat irritation, and congestion (Agency, 2011). Recent epidemiologic studies emphasize the relationship between Ozone and the mortality rate (Janke *et al.*, 2009) and asthma exacerbation (Currie & Neidell, 2004), (Laurent *et al.*, 2007), (Wilhelm *et al.*, 2009).

1.3. Presentation of the dataset

I use data on the concentration of pollutants and mortality rates available at a local level for all the whole of France.

Detailed data on atmospheric pollution come from the information system of the air quality measure (BDQA) used by the French Environment and Energy Management Agency (ADEME). ADEME gathers information coming from the 38 associations (AASQA) within the ATMO federation which monitor air quality. A large number of monitoring stations make up the federation. The French nomenclature identifies seven classes of stations, consistent with the various classifications defined at the European level : roadside, urban, industrial, near city background, national rural, regional rural, specific observations numbering 84, 286, 119, 138, 10, 62, and 13 respectively. Most of the monitoring stations are placed where the density of population is significant, apart from national rural monitoring stations. The measure taken into consideration in the study is the annual mean of concentration for pollutants within a civil year (1st January to 31st December) calculated by each AASQA for each captor and measured in micrograms per cubic meter of air. In principle the more disaggregated data is more desirable to cope with ecological inferences, but the health authority estimates are based on surveys with relatively small samples and are therefore less reliable. However, aggregate data may offer valuable clues about individual behavior. I divide the dataset in subsamples as an attempt to deal with the problems of confounding and aggregation bias. This annual mean is calculated by the ASQAA from the hourly mean for each monitoring station. This unit of concentration is mostly used to monitor outdoor air quality. Air pollutant concentrations do not necessarily produce accurate predictions of exposure levels. People may be resident in one area, but work in another. Nevertheless, the geographical level used in this article reduces the bias related to population mobility. The department surface represents an average of 570 000 hectares and we know from INSEE data that the average distance between the place of residence and the place of work is nearly 20km, so the accuracy of the exposure levels seems reasonable.

For spatial interpolation between monitoring stations, I use a geostatistical method that takes into account spatial dependence. This method does not necessarly reduce the amount of measurement error in the variable. The extent of measurement error is going to be greater for those departments with few monitoring stations where the population is more dispersed or lower. Lower or higher levels of the dependent variable within departments also induce measurement error. For example, more rural areas tend to be more agriculturally based and this may have an impact on mortality rates. Nevertheless, measurement error, even if not systematic, can induce attenuation bias.

Following Currie & Neidell (Currie & Neidell, 2004), I assign annual pollutant concentrations to the 95 French departments. Using the geographical coordinates of the census blocks with the highest population density of a local authority, I calculate the distance between the census blocks with the highest population density and all monitoring stations as it is explained in Appendix 1. First, I calculate the centroid of each local authority. I then measure the distance between the monitoring station and the center of the local authority. This distance corresponds to the weight attributed to a monitoring station, using the inverse of the distance to the center of the local authority. In order to assess the accuracy of our measure, I compare the actual level of pollution at each monitor location with the level of pollution that I would assign using the method previously described. The correlations between the actual and predicted levels of pollution are quite high for O_3, NO_2 and PM_{10} (0.6, 0.85 and 0.7 respectively) suggesting that the measure is quite accurate.

Table 1.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Pollutant variables					
NO2μg/m^3	31.714	11.375	12	74.046	220
O μg/m^3	53.029	15.39	30.601	99.653	220
PM10μg/m^3	21.461	5.508	8.818	57.384	220
Mortality rates					
Overall mortality rate	819.759	64.615	620	1000	220
Ms (%)	626.536	47.623	499	756	220
Mr (%)	1092.595	96.551	792	1393	220
Socioeconomics variables					
Income (%)	15303.625	2703.677	11011.659	27079.313	220
Un (%)	8.436	2.033	4.575	14.625	220
Education (%)	15.678	5.045	10.241	37.481	220
Poverty gap	0.5	0.501	0	1	220
Weather variables					
Sun (hours)	1979.783	362.492	1367.4	2962.3	205
Pr (mm)	2.211	0.571	0.855	3.865	214
Wind (km/hour)	100.212	13.056	68.400	147.6	208
Frost (days)	41.61	22.186	4	114	213
Demographics variables					
Sm (%)	1218.513	274.182	483.5	2298.7	220
Pop	771461.595	539704.805	123561	2561038	220
Industry (%)	16.175	4.778	5.93	24.476	220
PPHB (%)	136.174	37.84	71.696	268.82	220
Alcohol (%)	262.614	223.659	60	1594	220
Accident (%)	12.055	4.107	3	22	220
Atmo Index					
Atmo index 8 to 10	3.524	5.532	0	28	220

The top panel of Table 1.1 presents descriptive statistics for pollution data. NO,

NO_2 and PM_{10} are positively correlated with correlation coefficients between 0.5 and 0.8 as we can see in Table 1.2. They are negatively correlated with O_3 which may be due to the fact that Ozone is rapidly destroyed to form NO_2 within cities the correlation between both NO_2 and NO is high (0.85), so that I choose to keep NO_2 as an explanatory variable and drop NO to prevent autocorrelation. Moreover, I do not include observations for SO_2 and CO, as few monitoring stations measure these pollutants.

Table 1.2: pairwise correlation coefficients with significance level

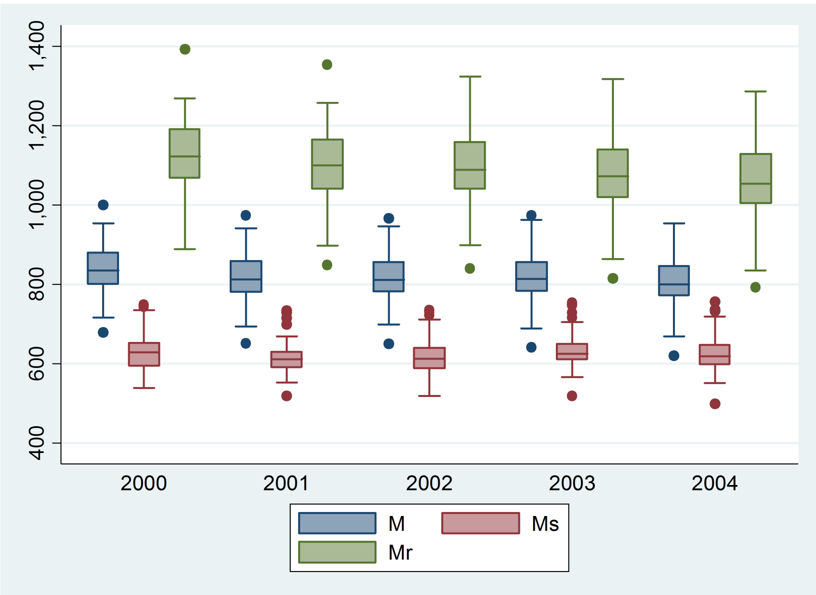
	NO2	PM10	O	NO	Atmo index(8-10)	Temperature
NO2	1.0000					
PM10	0.6086	1.0000				
	(0.0000)					
O	-0.3238	0.0301	1.0000			
	(0.0000)	(0.6566)				
NO	0.9644	0.5921	-0.3265	1.0000		
	(0.0000)	(0.0000)	(0.0000)			
Atmo index(8-10)	0.2663	0.5121	0.3833	0.1473	1.0000	
	(0.0034)	(0.0000)	(0.0000)	(0.1099)		
Temperature	0.1818	0.3079	0.4762	0.1447	0.2593	1.0000
	(0.0077)	(0.0000)	(0.0000)	(0.0344)	(0.0044)	

Other pollutants are also likely to be associated with differences in mortality, but data were unavailable to perform intra urban interpolations for these pollutants. Note that the local authorities with missing air pollution measures are all less populated areas. It is important to stress that air pollutant concentrations used to be below the limit value fixed by European and national institutions above which health can be harmed. In France, the threshold for NO_2, fixed by the European act 2002-13 related to air quality, is 200 $\mu g/m^3$ over 24 hours to protect human health. For a long term exposition (over one year) the regulated levels is $40 \mu g/m^3$ with respect to World health organization. The maximum concentration of NO_2 presented in the dataset is over this threshold. However, the annual mean for NO_2 is below the regulated level. The annual mean is the measure I use in my estimations. It corresponds to 'very good' air quality according to the ATMO index presented in Table 9. Moreover, the average concentration from the measure is lower than the level used by studies in the United States and even lower than in England (Janke *et al.*, 2009), where the level at which it is considered to harm health is already quite low. However, the E.R.P.U.R.S project in France shows that NO_2 and PM_{10} have a negative impact on health, even at low air concentrations, considering hospitalization numbers as the explicative variable (Campagna *et al.*, 2003). Pascal & al. in France obtain similar results, considering also different mortality rates in nine polluted cities (Pascal *et al.*, 2009) .

The second panel of Table 1.1 presents non incidental mortality rates. I consider a period of 5 years (2000-2004). The year corresponds to mid-year of the triennial period used. A moving average makes it possible to "smooth" a series of values expressed according to time. It is used to eliminate the least significant fluctuations. Mortality rate is a moving average of order 3. Data on mortality are available from 1980 to 2004 whereas data on pollution only exist from 1985 to 2005 with very few values before 2000. A large range of pollutants are responsible for outdoor air pollution, so that it is difficult to assign them to a specific health effect. This is why I use an non incidental mortality rate. We do not include in this paper specific causes of mortality, due to the weak variability of these data in France for the 2000

- 2004 period which does not allow any estimations. For instance, transport release $N0_2$ such that a high level of pollution may be observed next to roads where road accidents occur. As I wish to extract the only effect of pollution on mortality, I work on non incidental rate. The data on health come from the National Federation of Regional Health Observatories (ORS). I use age-standardized rates to control for different age structures across departments [3]. The standard deviation is quite high, showing that the data are spread out over a large range of values. The degree of dispersion (spread) and skewness in the data are presented graphically in Figure 1.1.

Figure 1.1: The yearly distribution of all causes mortality rates for the 2000-2004 period, in all departments.



3. This age-standardized rate is calculated as follows: $\sum_{i=1}^{19} PiTi$. Pi represents the share of age group for the population of reference and Ti represents the specific rate of mortality observed within a department for the age group i.

The third panel shows the socioeconomic variables: Income, education, poverty gap and unemployment based on the 2007 census of INSEE and the French Ministry of Labour (DARES). Definitions of the variables are given in Table 1.3. Note that data about ethnicity or race do not exist in France. The French Institute of Statistics does not collect data about language, religion, or ethnicity on the principle of the secular and unitary nature of the French Republic.

Table 1.3: Definition of variables

Variable	Definition	Sources
M	Total mortality rate, age standardized rates 2000-2004 calculated using data on registered deaths from INSERM, CEPIDc and INSEE. The year corresponds to mid-year of the trienal period used. Unit: per 100 000 people	National federation of regional health observatories (ORS)
Mr	Total male mortality rate, age standardized rates 2000-2004 calculated using data on registered deaths from INSERM, CEPIDc and INSEE. The year corresponds to mid-year of the trienal period used. Unit: per 100 000 people	National federation of regional health observatories (ORS)
Ms	Total female mortality rate, age standardized rates 2000-2004 calculated using data on registered deaths from INSERM, CEPIDc and INSEE. The year corresponds to mid-year of the trienal period used. Unit: per 100 000 people	National federation of regional health observatories (ORS)
NO_2, PM_{10},O_3	Annual mean of NO_2, PM_{10}, O_3 concentration (μg/m^3) respectively 2000-2004	French Environment and Energy Management Agency (ADEME)
Pr	High precipitation totals 2000-2004	Météo France
Sun	Annual cumulation of insolation in hours 2000-2004	Météo France
Frost	Annual number of frost days in days 2000-2004	Météo France
Wind	Annual instantaneaous maximum wind in km/hour 2000-2004	Météo France
Sm	Number of cigarettes sold for 1000 residents 2000-2004	French Monitoring Centre for Drugs and Drug Addictions (OFTD)
Alcohol	Annual number of deaths related to alcohol. It includes liver cirrhosis, alcoholic psychosis and alcoholism, cancer of the upper aero-digestive tract	National federation of regional health observatories (ORS)
Accident	Road accident rate, age standardized rates 2000-2004 calculated using data on registered road accident. Unit: per 100 000 people.	National federation of regional health observatories (ORS)
PPHB	Number of people per 1 hospital bed 2000-2004	National federation of regional health observatories (ORS)
Industry	Share of industry in the total value added of a department (in %).	French National Institute for Statistics (INSEE), census 2005
Education	Population from 15 years (without students) with minimum BAC+2 divided by population within department in 2006	French National Institute for Statistics (INSEE), census 2007
Un	The unemployment rate is the percentage of unemployed people in the labour force (occupied labour force + the unemployed) 2000-2004.	French Ministry of Labour (DARES)
Income	Income is defined as the net taxable income divided by the number of tax households within a department	French National Institute for Statistics (INSEE)
The intensity of poverty	(or poverty gap) is an indicator used to assess the extent to which the standard of living of the poor population is under the poverty line. It is calculated formally as follows:(poverty threshold - median standard of living of the poor population) / poverty threshold.	French National Institute for Statistics (INSEE), census 2005

The following panel describes the control variables. Data on weather come from Meteo France through the French Institute of the Environment (IFEN). Smoking rate fell by 35% between 2000 and 2004, probably due to the "Loi Évin" of 1991 and the tax increase (INSEE). Road accident rate fells 29% according to the data from the National Federation of Regional Health Observatories (ORS). I also collect from the ORS the number of people per hospital bed to measure the health care system and the availability of medical care resources in a particular department from 2000 to 2004. I add the share of industry to control for industrialization,[4] as a time invariant variable for each department based on the 2005 census of the French Institute of Statistics (INSEE).

Finally, the last row of descriptive statistics corresponds to an air pollution index. To capture peaks of pollution, I use the ATMO index calculated by the AASQA. The Atmo outlook varies daily according to air quality using a scale of 1-10 (1 = very good air quality, 10 = very bad air quality). This index takes into consideration the concentration of four subindexes characterizing Nitrogen Dioxide (NO_2), Sulphur Dioxide (SO_2), Particles in suspension (PS) and Ozone (O_3). It considers pollution measured only by urban and industrial monitoring stations for main agglomerations for a period from 2000 to 2003. After 2003, the construction of the index was changed, so that I cannot consider it for 2004. I retain 41 agglomerations and I associate each one with a department. I construct a yearly variable summing up

4. French data about industrialization and GDP are not precise enough to take into account time fluctuations among departments from 2000 to 2004. pourcentage of industry added-value over the total-added value for each department is available only every five years (INSEE).

the number of days above indices 8, 9 and 10, which corresponds to poor air quality according to the definition of the Atmo index (Table 1.4).

Peaks of pollution are correlated positively with ambient air pollution which gives more credence to the measure. This index variable is positively correlated with the previous measure of NO_2, NO, PM_{10} and O_3. However, further in the estimation, I prefer to use real concentrations of pollution instead of indices. In fact, few days correspond to peaks of pollution, and fixing a threshold below which pollution does not have any impact is highly arguable. Pollution does indeed fluctuate, a low level can be active and the level perceived as toxic is variable, even among the healthy population. Within a population, some people are more sensitive than others and will suffer from atmospheric pollution even at really low levels ; levels below the actual threshold fixed by the public authorities. I aim to test this intuition.

Table 1.4: The Atmo index

Index scale	Subindexes	PM_{10} **scale** Average of mean daily concentrations in $\mu g/m^3$	NO_2 **scale** Average of the hourly maxima in $\mu g/m^3$	O_3 **scale**
Very good	1	0 - 9	0 - 39	0 - 29
Very good	2	10-19	40 - 79	30 - 54
Good	3	20 - 29	80 - 119	55 - 79
Good	4	30 - 39	120 - 159	80 - 104
Moderate	5	40 - 49	160 - 199	105 - 129
Poor	6	50 - 64	200 - 249	130 - 149
Poor	7	65 - 79	250 - 299	150 - 179
Bad	8	80 - 99	300 - 399	180 - 209
Bad	9	100 - 124	400 - 499	210 - 239
Very bad	10	125 and more	500 and more	240 and more

1.4. Model and Econometrics

1.4.1. Specification

The focus of this study is the relationship between average pollution, socioeconomic status, and mortality. the unit of analysis is the department, which is the main administrative unit below the national regional level. The department of France are French administrative divisions. The departments form one of the three levels of local government, together with the 22 metropolitan and 5 overseas regions above them. There are 95 departments in France with an average population of 620 000 people, ranging from over 70 000 to over two million. Departments are grouped within 22 metropolitan areas known as regions [5].

In the analysis, I start by estimating a standard model with the non incidental mortality rate as the explicative variable without considerations of environmental quality. After doing preliminary regressions for various functional forms and following the results from an overall normality test based on skewness and on kurtosis for each of them, I estimate an equation of the following form to ensure that errors are normally distributed $\varepsilon \sim N(0, \sigma^2)$ [6] :

$$X_{it}^{k} = \alpha_k + Socioeconomic_{it}\beta_k + Demographics_{it}\eta_k + Z_{it}\phi_k + \varepsilon_{it}^{k} \qquad (1.1)$$

5. Due to missing data, we remove departments from the analysis in order to consider a balanced panel. We end up with 41 departments

6. There is no evidence that the log transform is the best fit for mortality time trends (Bishai & Opuni, 2009). Moreover, given the size of the department, the effect of outliers may not be a problem here.

where i indexes the local authority, t indexes the year, k the kind of mortality rate. X_{it}^{k} is a vector of all causes mortality rates (overall mortality rate, male and female mortality rate). Socioeconomic variables and particularly the unemployment rate and income are included as the main explanatory variables. Due to multi-collinearity issues, I am not including both the average income and the education variable. The squared correlation between education and the average revenue is above 0.8.

The vector $Demographics_{it}$ includes several variables. First, it accounts for lifestyle, which refers to the regular activities and habits a person has that could have an effect on his or her health. I include the smoking rate variable as a proxy for lifestyle. The number of people per hospital bed $PPHB_{it}$ in each department is included as a proxy to measure the health care system and the availability of medical care resources in a particular department. I also include the % of industry added-value over the total added-value $Industry_i$ for each department as a time-invariant variable. I also take weather patterns into consideration at department level Z_{it} as a control for average pollution levels. I consider the annual mean of precipitation to capture the effect of very wet years, the maximum wind speed, the number of frost days, and the annual cumul of sunlight as a time varying control. Some studies contend that mostly long-term (i.e., monthly and annual) fluctuations in temperature affect mortality (Martens, 1998). Besides, wind speed measurements are important for air quality monitoring. The higher the wind speed, the lower the pollutant concentration. Wind dilutes pollutants and rapidly disperses them throughout the immediate area. ε_{it} is the error term.

Independent variables have explained most differences between departments and years, but there is probably some unmodeled heterogeneity. Thus, the next step in performing a multilevel analysis is to decide whether the explanatory variables considered in the analysis have fixed or random effects. The Hausman test considers the null hypothesis that the coefficients estimated by the efficient random effects estimator are the same as the ones estimated by the consistent fixed effects estimator. By running this test, the fixed effect model appears to be the most efficient one. In fact, I think of each department as having its own systematic baseline. I also calculate the robust variance estimator, in order to prevent the heteroskedasticity that I found by running Breush-Pagan test: this test checks if squared errors are explained by explanatory variables. the estimation will also take into account autocorrelation, because the Wooldridge test shows that disturbances exhibit autocorrelation, with the values in a given period depending on the values of the same series in previous periods. To address the possibility that omitted variables account for some of the heterogeneity among French departments, an error component model is estimated:

$$\varepsilon_{it} = c_i + \delta_t + u_{it} \tag{1.2}$$

c_i and δ_t are residual differences where c_i is a department effect which accounts for differences across departments that are time-invariant (e.g lifestyle differences that we cannot take into account), δ_t is a year effect which controls for factors that vary uniformly across departments over time, and u_{it} is the remaining error term [7].

It is also likely that a population's health affects unemployment via productivity, education and other factors. This potential simultaneity can be a source of endogeneity, making standard estimators inconsistent I need to test this hypothesis, so I consider the lag of the endogenous variable, unemployment, as an instrument. The F-test on the excluded instruments in the first stage regression confirms the validity of the instrument. To avoid the weak instruments pathology, we look at the F-test on the excluded instruments in the first stage regression and check whether the test statistic is greater than 10 (F(1,192) = 28.91). Then, the Hausman test rejects the endogeneity of the model (P=0.810).

This paper is also concerned with spatial correlations which would bias the results or introduce inefficiency. If the observations are spatially clustered, the estimates obtained will be biased or inefficiency will be introduced. In fact, the mortality rate in one region could be related to that in another. [8]. As a consequence, I will calculate the Driscoll and Kraay non-parametric adjustment of standard errors model allowing for both space and time adjustments.

In the second model, the mortality rate is expressed as a function of environmen-

7. The Ramsey test confirms the robustness of the specification.

8. The Moran Index of spatial contiguity rejects the null hypothesis that there is no spatial clustering of the value in the raw mortality data. First tail test: I=0.266 at a 1 pourcent probability

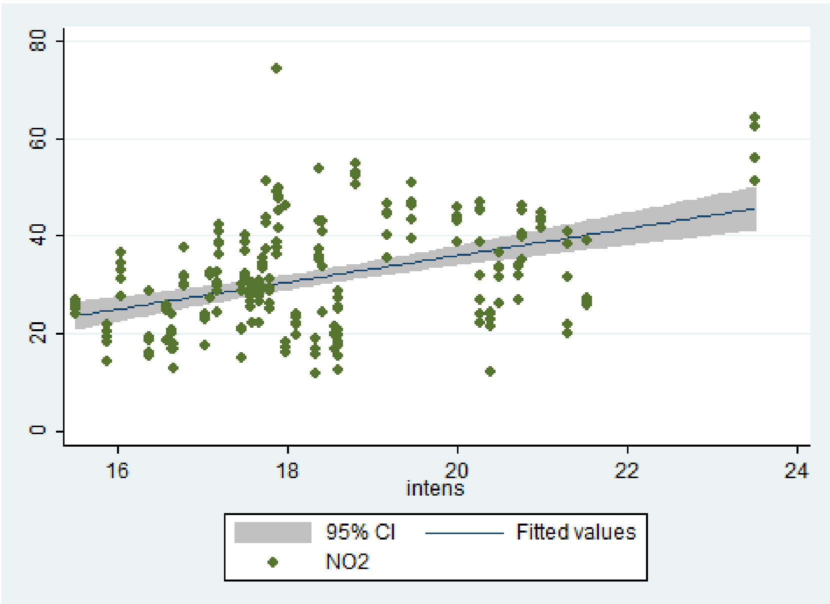
tal variables added to the previous variables. I will estimate the model :

$$X_{it}^k = \lambda_k + P_{it}\theta_k + Socioeconomic_{it}\psi_k + Demographics_{it}\sigma_k + Z_{it}\phi_k + \varepsilon_{it}^k \quad (1.3)$$

P_{it} is a vector of air pollutant concentrations for O_3, NO_2 and PM_{10}. In this model, the main coefficient of interest is θ representing the mean parameter estimates for all 95 departments for explanatory variables P_{it}. It also represents the effect of air quality on health outcomes. I will again use the fixed effect estimator to control for heterogeneity between departments, with and without the Driscoll and Kraay standard errors model. Besides, the endogeneity problem has to be discussed in this context. I include fixed effects and some controls to address the problem of unobservable variables. However; there may be time-varying unobservable variables, not common to all regions and not captured by the dummies, which could bias the estimates. One may argue that European unemployment fluctuates around a very low level (Blanchard, 1986) making the previous Hausman test really weak. An association between the business cycle and mortality could, for instance, be driving the result (Chay & Greenstone, 2003). However, the French Statistical Institute does have access to the yearly business cycle data for each department.Another endogeneity bias could be that people may move in response to pollution levels. People who care more about health, and hence live a healthier life style, may move to less polluted areas, introducing an upper bias in the estimate of pollutants. However, migrations in France between departments are essentially in border areas and are mainly due to preferences for urbanization (INSEE). .

Finally, I estimate a model dividing my sample with those above and below the median of poverty gap. To do so, I create a dummy with respect to the intensity of poverty. 50 % of departments are below and 50% are above this median. I want to study whether the impact of pollution on health is greater when I consider poorer population. Figure 1.3 shows a potential positive relationship between NO_2 and the intensity of poverty.

Figure 1.2: Correlation between poverty gap and NO_2



Departments with a high pollution level seems to be the one with a low socioeconomic level. Besides, I ask oneself if a change in pollution benefits, in term of health, even more to high socioeconomic than low levels areas. The main purpose is to figure out if inequalities tend to mount up within French department. People with low incomes may be disproportionately vulnerable as well as disproportionately exposed. Is the health damages of an increase in air pollution bigger in poorer area compared to its counterparts? I show in the next section that socioeconomic factors, in particular the unemployment rate, greatly interfer when studying the impact of NO_2 on mortality rate at different level of poverty.

1.4.2. Results

1.4.2.1. Impact of environment quality on health

I start by examining a standard model of mortality rate without consideration of environmental quality. I then add NO_2, O_3 and PM_{10} to the specification to see if considering pollutant variables improves the global fit of the model. To capture the department effect and the spatial autocorrelation, both fixed effects clustered at the regional level and the Driscoll and Kraay standard errors with fixed effect are estimated. Approximately seventy percent of the variation in the response variable may be attributed to explanatory variables.

I first estimate the standard model with the OLS, trying to test the most complete model, and I observe in the first column from Table 1.5 that all the coefficients of the determinants of mortality are significant. However, I also use the within estimator as I assume that the unobserved factors f_{it} between departments determine both mortality rates and explanatory variables. I observe a loss of significativity for some coefficients which may be due to the correlation between department-specific effect and both explanatory variables. The Fixed effect imposes time-independent

Table 1.5: Standard model of non incidental mortality rate

	OLS	FE	D-K S.E
Income	-0.0165***	-0.0641***	-0.0641**
	(0.00253)	(0.00596)	(0.0141)
Un	8.217*	17.01**	17.01***
	(4.435)	(7.432)	(2.593)
Pr	-16.83	3.755	3.755
	(9.914)	(5.898)	(5.744)
Sun	-0.0793***	0.0764***	0.0764**
	(0.0178)	(0.0174)	(0.0206)
Wind	0.786***	0.271**	0.271
	(0.267)	(0.120)	(0.199)
Frost	0.728**	-0.966***	-0.966*
	(0.290)	(0.209)	(0.425)
Sm	0.0619**	-0.105***	-0.105*
	(0.0231)	(0.0247)	(0.0412)
Industry			
PPHB	0.0317	1.690	1.690
	(0.237)	(2.411)	(0.876)
Alcohol	0.0840	0.174*	0.174**
	(0.0508)	(0.0901)	(0.0613)
department FE		x	x
Observations	203	203	203
R-squared	0.598	0.753	

[a]

[a] Notes: This table presents the standard model of non incidental mortality rate with its main determinants. All regressions are estimated with standard errors clustered at the regional level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

effects for each entity that are possibly correlated with the regressors, which is why $Industry_i$, a time invariant variable, is not taken into account. The last column shows the Driscoll and Kraay standard errors model which takes spatial autocorrelation into account, with fixed effects. I observe that income impacts negatively mortality rate in every regressions at a 1% level of significativity. I am in line with the literature saying that income is a significant determinant of health. Ettner finds that increases in income significantly improve mental and physical health (Ettner, 1996). Inadequate education and living conditions ranging from low income to the unhealthy characteristics of neighborhoods and communities can harm health through complex pathways. Health disparities by income is partly explained by disparities in medical care. French departments with a high level of income are more likely to have a low mortality rate than department with a lower income. As a consequence, a shortsighted political focus on reducing spending in education, child care, jobs, community and economic revitalization, housing, transportation, could actually increase medical costs by magnifying disease burden and widening health disparities.

I then study the relationship between NO_2, O_3, PM_{10} and mortality rates in both a single pollutant model (Table 1.6) and in a multi-pollutant one (Table 1.7). The multi-pollutant model allows coefficients to be examined at the same time, so as to not overestimate the impact of one pollutant.

As shown in the single pollutant model, coefficients for PM_{10} are significant with the fixed effect model clustered at the regional level and the Driscoll and Kraay estimation. However, coefficients are not significantly different from zero when I consider a multiple pollutant model in both specifications. The variation I have in the dataset may not be sufficient to obtain significant results for PM_{10} even though the single pollutant model shows a significant impact of PM_{10} on mortality rates. Mortality rates do not vary too much within region over time. moreover, we only have access to the variation within departments over a few years, from 2000 to 2004. This may be very little variation, perhaps some of it due to measurement error which would bias coefficients towards zero. It also may be explained by the interaction be-

Table 1.6: A simple pollutant model of mortality

VARIABLES	(1) FE	(2) D-K S.E	(3) FE	(4) D-K S.E	(5) FE	(6) D-K S.E
NO2	0.438*	0.438*				
	(0.237)	(0.182)				
O			0.220	0.220		
			(0.377)	(0.195)		
PM10					0.193	0.193*
					(0.263)	(0.0881)
Income	-0.0645***	-0.0645**	-0.0646***	-0.0646***	-0.0649***	-0.0649***
	(0.00568)	(0.0142)	(0.00543)	(0.0134)	(0.00560)	(0.0134)
department FE	x	x	x	x	x	x
Weather controls	x	x	x	x	x	x
Demographic controls	x	x	x	x	x	x
Socioeconomic controls	x	x	x	x	x	x
Observations	203	203	203	203	203	203
Adjusted R-squared	0.756		0.754		0.754	

[a]

a. Notes: This table presents the impact of NO_2, O_3 and PM_{10} on non incidental mortality rate. All regressions are estimated using fixed effect with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather pattern with wind, sun, precipitations and frost; a vector of socioeconomic variables including unemployment rate and income; and a vector of demographics including the level of indutrialization, people per hospital bed and the smoking rate. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

tween O_3 and NOx which may biased the result obtained for PM_{10}. This result is in line with Chay et al. who examine the effect of particulate matter on adult mortality in the US during the 1970s. They find no impact of this source of pollution on adult mortality (Chay *et al.*, 2003). In contrast, this result is opposed to the French study by the Sanitary Health Institute which found a positive effect of PM_{10} on mortality in a panel of nine different French cities (Pascal *et al.*, 2009). However, this study does not precise the type of estimator used. Furthermore, Pascal et al. do not take into account the influence of lifestyle or socioeconomic factors on health as their model strictly includes weather data whereas the robustness of the model is not verified if I take socioeconomic factors out. Finally, the average concentration from the measure is probably lower than the level used by the Sanitary Health Institute which considers 9 urban cities.

Ozone is negatively correlated with the mortality rate but it is not significant in both a simple and multiple pollutant model. The relationship between Ozone and temperature remains a complex phenomenon and may be the cause of the negative coefficient also pointed out in England (Janke *et al.*, 2009). It seems highy complex to extract one effect from the other. The Sanitary Health Institute in France emphasizes the complexity of studying the interaction between Ozone and sanitary variables because temperature, humidity, winds, and the presence of other chemicals in the atmosphere influence Ozone formation, and the presence of Ozone, in turn, affects those atmospheric constituents. French data show this positive correlation between temperature and Ozone.

In contrast, NO_2 appears to have a significant and positive effect in both single and multi-pollutant models when I consider the fixed effect regression model with Driscoll and Kraay standard errors. These results suggest NO_2 has a positive a significant impact on mortality rate in France at a department level. The fixed effect estimate suggests that, per 3 $\mu g/m^3$ increase in NO_2, there is almost one more death a year per 100.000 [9]. Concentrations of NO_2 vary from 12 to 74 $\mu g/m^3$, suggesting a difference of nearly 20 deaths a year per 100.000 depending on the department. This result confirms the existence at a high level of pollution of a

9. The death are registered in the municipality of death

long-term relationship between current air pollution levels and mortality in France [10]. I also observe from the OLS estimator that the effect of NO2 on the mortality rate tends to lead to erroneous conclusion if the fixed-effects problems are neglected. As a result, I give more credence to fixed effect estimators and specially the fixed effect regression model with Driscoll and Kraay standard errors for the rest of the study. As the multiple pollutant model does not show any significativity for PM_{10} and Ozone for the reasons I explained previously, I focus on the unique pollutant, NO_2, as it has a relevant significativity in both models. To deal with the strong correlation existing between the three pollutants, I replace pollution variables by the atmo index in the second block of the multiple pollutant model of mortality. However, the atmo index variable is not significative in any of the three models estimated. Atmo index may be not precise enough to highlight a pollution effect.

10. Short-term effects studies refers to daily variation whereas long-term studies use cohort studies over several years

Table 1.7: A multiple pollutant model of mortality

VARIABLES	(1) OLS	(2) FE	(3) D-K S.E	(4) OLS	(5) FE	(6) D-K S.E
NO2	0.261	0.611*	0.611**			
	(0.596)	(0.359)	(0.180)			
PM10	0.0990	-0.285	-0.285*			
	(0.505)	(0.424)	(0.124)			
O	-0.169	0.272	0.272			
	(0.535)	(0.337)	(0.167)			
atmo				-0.665	-0.542	-0.542
				(1.104)	(0.460)	(0.330)
Income	-0.0165***	-0.0645***	-0.0645**	-0.0215***	-0.0891***	-0.0891***
	(0.00218)	(0.00560)	(0.0143)	(0.00368)	(0.0103)	(0.0103)
Constant	901.8***	1,518***	0	995.1***	1,694***	0
	(121.3)	(254.1)	(0)	(152.0)	(319.5)	(0)
department FE		x	x		x	x
Weather controls	x	x	x	x	x	x
Demographic controls	x	x	x	x	x	x
Socioeconomic controls	x	x	x	x	x	x
Observations	203	203	203	119	119	119
Adjusted R-squared	0.745	0.581		0.619	0.811	

[a]

a. Notes: This table presents the impact of a multiple pollutant model on non incidental mortality rate . All regressions are estimated using fixed effect with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather pattern with wind, sun, precipitations and frost; a vector of socioeconomic variables including unemployment rate and income; and a vector of demographics including the level of indutrialization, people per hospital bed and the smoking rate. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

1.4.2.2. Gender Analysis

I now consider separately female and male mortality rates as explicative variables related to NO_2. Results are detailed in Table 1.8. Income is again negatively and highy correlated with female and male mortality rate for both models. I observe a significant effect for females as I do not find any for males. I am in line with the previous literature (Clougherty, 2010) (Marr, 2010): women are more sensitive and vulnerable to pollution. Smoking rates and unemployment have a significant and positive effect on male mortality rates when considering fixed effect models. In contrast, the female mortality rate model shows significativity for the pollution variable but not for the smoking rate. The female fixed effect estimate suggests that, per 2 μg/m^3 increase in NO_2, another death for women is registered, suggesting a difference of 30 deaths a year for women across departments. Lifestyle represented here with the smoking rate seems to be more significant than air pollution concentrations for male mortality rates. moreover, The relative impact of having a job on health seems to be greater for men. This result leads us to think about the significance of considering the individual degree of exposure including demographics, type of activities or personal health situation.

Table 1.8: A gender model of mortality

	female Mortality rate		male Mortality rate	
VARIABLES	FE	D-K S.E	FE	D-K S.E
NO2	0.566***	0.566***	0.00910	0.00910
	(0.181)	(0.0760)	(0.242)	(0.145)
Income	-0.00540*	-0.00540	-0.0254***	-0.0254***
	(0.00294)	(0.00284)	(0.00491)	(0.00322)
Sm	-0.0114	-0.0114	0.0135	0.0135
	(0.0120)	(0.00873)	(0.0140)	(0.00635)
PPHB	0.471	0.471	1.598	1.598**
	(1.396)	(0.514)	(2.765)	(0.498)
Un	0.881	0.881	4.654	4.654***
	(3.103)	(0.669)	(2.905)	(0.494)
department FE	x	x	x	x
Weather controls	x	x	x	x
Demographic controls	x	x	x	x
Socioeconomic controls	x	x	x	x
Observations	203	203	203	203
Adjusted R-squared	0.402		0.764	

[a]

a. Notes: This table presents the impact of a multiple pollutant model on non incidental mortality rate . All regressions are estimated using fixed effect with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather pattern with wind, sun, precipitations and frost. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

1.4.2.3. Interaction between socioeconomic status and environment quality

Furthermore, I might suspect that exposure to air pollution related to health varies with socio-economic status in France. People with low incomes may be disproportionately exposed to environmental contamination that threatens their health.

First, the previous models of this paper shows that income is negatively and very significatively related to mortality rate. Richer departments have a lower mortality rate. However, this study does not answer precisely the reasons behind this result. Better access to health care may be a reason why mortality rates in richer geographical areas are lower. The variable people per hospital bed is significant and positive in the OLS estimation in Table 1 giving credibility to this intuition. moreover, there is a positive correlation between the level of income and people per hospital bed in the dataset [11].

If we think about environmental justice, the spatial distribution of pollution among department may also be another reason why mortality rates in richer geographical areas are lower. Lower pollution in richer department may influence positively their health relatively to departments with a high intensity of poverty. Table 1.9, which present average values of mortality rates above and below the median of income supports this idea.

Secondly, poor people may also be more sensitive to pollution which threatens their health. To study this possibility, I divide the panel into departments above

11. The simple coefficient of correlation indicates a positive correlation of 0.2 between income and people per hospital bed within the French departments

Table 1.9: the average values of mortality rates for different thresholds of pollution and socioeconomic status

Variable	Mean	Std. Dev.	Min.	Max.	N
	Low income(below the median)				
Non incidental mortality rate	781.8545	67.2788	670	1000	110
	High Income (above the median)				
Non incidental mortality rate	745.7455	68.35437	584	915	110
	Low pollution(below the median)				
Non incidental mortality rate	773.9455	62.39896	670	953	112
	High pollution (above the median)				
Non incidental mortality rate	753.6545	75.86364	584	1000	108

and those below the median of the intensity of poverty as shown in Table 1.10. The intensity of poverty is an indicator used by INSEE to assess the extent to which the standard of living of the poor population is under the poverty line. The higher the indicator, the greater the poverty gap is said to be, in that the standard of living of the poorest is a very long way below the poverty threshold.

The first two columns of Table 1.10 presents the result for the sample above the median of poverty gap compared to the last two columns which represents the sample below. I observe that the coefficient for income remains highy significant and negative in both sample suggesting again the existence of health disparities with respect to income. The size of the impact is similar for both samples. The level of income does not have a higher impact on mortality rate when considering departments with a high poverty gap with respect to its counterparts. We cannot conclude to a poverty trap effect from this result. I observe a similar result for the impact of NO_2 concentration on mortality rate. Besides, I note that the sample with a low intensity of poverty seems slighly more affected by pollution than the sample above the median of poverty gap. This result is opposed to the intuition [12]. Coefficients difference for NO_2 may come from the unemployment variable. Unemployment seems to moderate the effect of NO2 on mortality rate. The coefficient for unemployment is indeed positive and significant in the sample above the median of the intensity of poverty whereas it is not significant within department with a lower poverty gap. This result suggests the impact of unemployment on mortality rate relatively to pollution is higher for department with a high poverty gap than for department with a low poverty gap. Unemployment tends to reduce the effect of NO_2 on mortality rates. Poorer a department is, higher the effect of unemployment relatively to pollution is. The variable of interaction between NO_2 and unemployment presented in Table 1.11 confirms this intuition.

Table 1.10: A model with respect to poverty gap

	(1)	(2)	(3)	(4)
	high poverty gap		low poverty gap	
VARIABLES	FE	D-K S.E	FE	D-K S.E
NO2	0.165	0.165***	0.387**	0.387*
	(0.120)	(0.0319)	(0.176)	(0.144)
Income	-0.0134***	-0.0134***	-0.0117**	-0.0117*
	(0.00348)	(0.00158)	(0.00454)	(0.00466)
Un	4.725*	4.725	1.671	1.671
	(2.578)	(2.279)	(5.463)	(1.093)
Department FE	x	x	x	x
Weather controls	x	x	x	x
Demographic controls	x	x	x	x
Observations	94	94	109	109
Adjusted R-squared	0.629		0.444	

a

a. Notes: This table presents the impact of a multiple pollutant model on all causes mortality rate with respect to the level of poverty . All regressions are estimated using fixed effect with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather pattern with wind, sun, precipitations and frost; and a vector of demographics including people per hospital bed and the smoking rate. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

12. However, confidence intervals overlap with eachother suggesting both coefficients are not significantly different.

1.4.2.4. Robustness check

I perform several robustness tests to make sure the estimation is not biaised by any unobserved factors. In all the robustness estimations summarized in table 1.11, I take spatial autocorrelation and department effect into account, using the Driscoll and Kraay standard errors model with fixed effect.

First, to be more precise, I want to analyze whether the effect of the socioeconomic variables is moderated or modified by the introduction of the environmental variable. To do so, I include an interaction variable to look at how unemployment and NO_2 interact. I add an interactive term $P_{it}Socioeconomic_{it}$ between unemployment and environmental quality to provide a better description of the relationship between mortality rate and the independent variables such that:

$$X_{it}^k = \lambda_k + P_{it}\theta_k + Un_{it}\psi_k + P_{it}Un_{it}\varpi_k + Demographics_{it}\sigma_k + Z_{it}\phi_k + \varepsilon_{it}^k \quad (1.4)$$

$$= \lambda_k + (\theta_k + Un_{it}\varpi_k)P_{it} + Un_{it}\psi_k + Demographics_{it}\sigma_k + Z_{it}\phi_k + \varepsilon_{it}^k$$

where $(\theta_k + Un_{it}\varpi_k)$ represents the effect of environmental quality on the mortality rate at a specific level of socioeconomic variables and ϖ_k indicates how much the slope of P_{it} changes as the unemployment variable goes up or down by one unit. To ease the interpretation, I consider a dummy variable for NO_2. NO_2 will take the value of "1" when departments are above the median of NO_2 concentration and "0" otherwise[13]. The coefficients for unemployment, NO_2 and the interactive variable are significant with the endogenous variable "all causes mortality rate". The significativity of the interaction coefficient suggests that the effect of NO_2 has been modified by the unemployment variable. In other words, the effect of NO_2 $(\theta_k + Socioeconomic_{it}\varpi_k)$ at some value of unemployment $Socioeconomic_{it}$ has a significant effect on the mortality rate. And the negative sign indicates the introduction of unemployment moderates the effect of pollution on mortality which confirms the intuition described above.

Besides, interacting income and pollution does not show any significativity. Second, to deal with potential non linearity in the model, income squared is added to the estimation. The really small coefficient is not surprising knowing the average of 234.000.000. It may explain why the impact of income on mortality rate is not changing in the sample above and below the intensity of poverty. It also may be why there is no significativity when interacting pollution and income. Third, I include education to the estimation to make sure this variable does not biaised results. There is also, most likely, a direct positive effect of education on health (Groot & Maassen van den Brink, 2007). While the exact mechanism underlying this link is unclear, the differential use of health knowledge and technology is almost certainly an important part of the explanation. We cannot conclude on the sign of the coefficient for education because there is a high correlation between income and

13. Critics assert that an increased level of collinearity in models including a multiplicative term distorts the beta coefficients. However, a fixed effect model, or a mean purged regression model, automatically reduces multicollinearity.

education. However, the impact of NO_2 on mortality rate does not change neither in magnitude and in significativity when adding education to the model. Fourth, the introduction of population within the regression shows that population among department is not an issue. Fifth, I include a variable of pollution squared in order to compare results when facing higher atmospheric pollution. The positive and highy significant coefficient for NO_2 squared indicates that the impact of NO_2 on the overall mortality rate is amplified with higher pollution level. Finally, all over the paper, I have considered NO_2 as a measure of exposure to air pollution which was calculating using the annual measure of concentration weighted by the inverse of the distance between the census block where there is the most significant population density and the monitoring station. As a last robustness check, I estimate a model considering the concentration of NO2 non weighted. Model 8 shows it does not change the result if I consider a measure without any weight. The impact of NO_2 is still positive and significant [14].

14. However I prefer the weighted measure of concentration as I have described in the previous section because it tends to lead to a more accurate measure of exposure to air pollution.

Table 1.11: Robustness check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NO2	0.263**	0.966	1.487**	0.277***	0.340***			
	(0.0749)	(0.933)	(0.397)	(0.0600)	(0.0729)			
$Income^2$	-1.01e-06**							
	(2.66e-07)							
NO2*Income		-4.17e-05						
		(6.53e-05)						
NO2*Un			-0.122**					
			(0.0315)					
Education				46.23***				
				(7.713)				
Pop					-0.000329**			
					(0.000113)			
$NO2^2$						0.00516**		
						(0.00116)		
Non weighted NO2							0.649**	
							(0.191)	
department FE	x	x	x	x	x	x	x	x
Weather controls	x	x	x	x	x	x	x	x
Demographic controls	x	x	x	x	x	x	x	x
Socioeconomic controls	x	x	x	x	x	x	x	x
Observations	99	203	203	203	203	203	203	203

[a]

a. Notes: This table presents the impact of a multiple pollutant model on all causes mortality rate with respect to the level of poverty. All regressions are estimated using fixed effect and with Driscoll and Kraay standard errors, a vector of weather control including wind, sun, precipitations and frost, a vector of socioeconomic variable including unemployment rate and income and a vector demographics including people per hospital bed and and the smoking rate. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

These results does not show greater effects among the more deprived. Thus, I cannot conclude to the existence of environmental injustice among French departments. Health disparities exist among French departments but seem to be more related to socioeconomic factors than differences in sensitivity to pollution. In fact, high level of unemployment tends to moderate the effect of pollution on health. However, the impact of pollution on mortality rates remain an important issue as we found a significant effect of NO_2 especially at levels below the current standard.

1.5. Conclusion

The objective of this paper has been first to investigate whether a department's environmental quality and socioeconomic status relative to its neighbors has an impact on its mortality rate. The second purpose has been to analyze the link between inequalities and air quality across departments. I test these hypotheses by using a multivariate model and taking spatial autocorrelation and fixed effects into account.

The results are strongly supportive of the hypothesis that NO_2 has a positive impact on mortality with the effect being larger when considering higher level of pollution. Moreover I show that even relatively low concentrations of air pollutants, at levels below the regulated threshold, are related to a range of adverse health effects.

I also shed light on the existence of health disparities in France. As a consequence, the choice of economic policies can have severe implications for health and medical spending. I also find that there is a significant link between the unemployment rate and the concentration of NO_2 in a department. The impact of unemployment relatively to pollution is prevalent in poor areas compared to richer ones. Finally, I point out that women's health is more impacted than men's health by NO_2. This finding is consistent with the results of international studies that have examined the relationship between economic inequality, environmental quality and health. It also confirms the importance of ambient air pollution and reinforces the need for politicians to take into account environmental justice in France.

The paper suggests that further research on environmental inequality in France focusing on smaller geographical levels and individual characteristics is essential. It would be consistent to examine the impact of atmospheric pollution focusing on individual-level data. It would also be interesting to have access to morbidity data, especially for occupation-linked diseases, to shed light on the implications for loss of productivity.

1.6. Appendix

I assign annual pollutant concentrations to the 95 French departments. To do so, I assign a weight to every monitoring station. This weight corresponds to the distance between the census block of the department with the highest population density and the monitoring station. This distance is given as the great-circle distance between the two points, that is, the shortest distance over the earth's surface giving an 'as-the-crow-flies' distance. Let (L_i,N_i) be the latitude and longitude in degrees of monitoring station i and (Lj, Nj) of the census block with the highest population density j. The distance between the monitoring station i and the census block j is given by:

$$d_{ij} = arccos(G_{ij})R$$

where R is the radius of the earth, measured around the equator $(R = 6378)$ and

$$G_{ij} = sin(aL_i)sin(aL_j) + cos(aL_i)cos(aL_j)cos(aN_j - aN_i)$$

with $a = \pi/180$ From this distance I calculate a weighted mean of pollutant concentration. The weighting attributed to a monitoring station corresponds to the inverse of the distance between the the census block with the highest population density and the station so that every element C_{ij} of the distance matrix C is given by:

$$C_{ij}=\frac{1/d_{ij}}{\sum_{i=1}^{n} 1/d_{ij}}$$

Matrix C is a stochastic matrix of size N×N where elements in the main diagonal are zero. It is normalized in order to have each row summing to 1. Such normalization allows us to consider the relative distance instead of the absolute one. Then, I calculate the average weighted mean of pollutant concentration $\bar{P}$ within the entire department:

$$\bar{P}=\sum_{i=1}^{n} C_{ij} P_{iu}$$

P_{iu} corresponds to the annual mean concentration measured by the monitoring station i for the pollutant u.

Chapter 2

Energy Production and Health Externalities: Evidence from Oil Strike Refineries in France.[1]

1. This chapter is a joint work with Matthew Neidell from Columbia University

This paper examines the effect of energy production on newborn health using a recent strike that affected oil refineries in France as a natural experiment. First, we show that the temporary reduction in refining lead to a significant reduction in sulfur dioxide (SO_2) concentrations. Second, this shock significantly increased birth weight and gestational age of newborns, particularly for those exposed to the strike during the third trimester of pregnancy. Back-of-the-envelope calculations suggest that a 1 unit decline in SO_2 leads to a 196 million euro increase in lifetime earnings per birth cohort. This externality from oil refineries should be an important part of policy discussions surrounding the production of energy.

2.1. Introduction

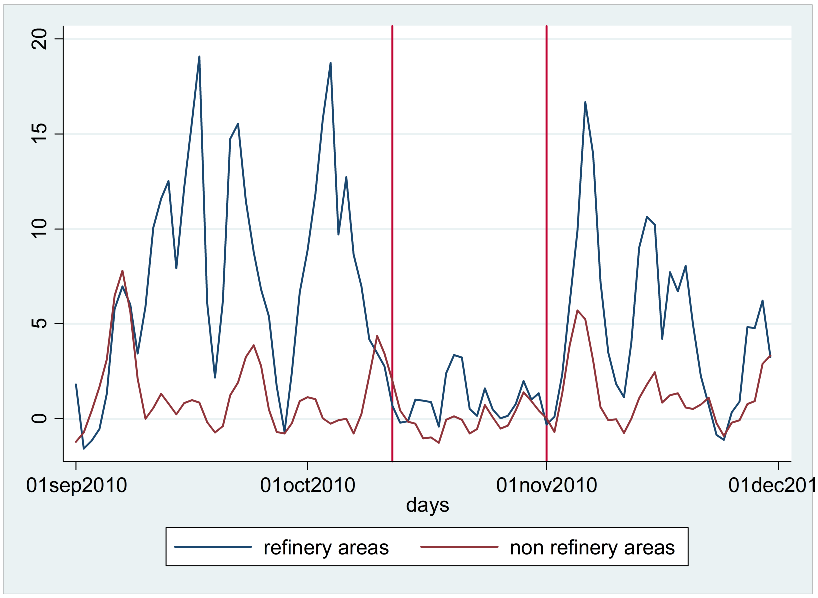
Meeting the continued increased demand for energy is a major issue faced by nearly all countries. While there is much interest in developing renewable sources of energy, oil remains the predominant source given its relative price. Its portability also makes it particularly attractive for mobile sources, suggesting a reprieve in energy demand is unlikely in light of the tremendous growth in automobile ownership and travel throughout the world. Despite the price advantage of oil, its production poses a health risk. The point source emissions include several pollutants linked with numerous health impacts, most notably sulfur dioxide (SO_2). In some countries, such as France, nearly 20 percent of ambient SO_2 emissions come from oil production (Soleille, 2004). Evidence links SO_2 with a wide range of respiratory effects, and as such is regulated under environmental policies throughout the world. The optimal design of energy policy must consider this production externality when comparing its full costs to those from renewable energy production. In this paper, we estimate the health effects from oil production by exploiting the pension reform strikes in France in October, 2010 that lead to a major disruption in the production of oil. These strikes provide an ideal natural experiment for overcoming the typical biases that arise when estimating the health effects of pollution. Amid nationwide protests over pension reform that involved raising the retirement age, striking workers blocked fuel supplies to oil refineries, which resulted in a complete cessation of operations at several major refineries for nearly a month. As Figure 1 demonstrates, this lead to a sharp reduction in SO_2 in areas close to the refineries when the strike began, that quickly dissipated once the strike was resolved and production resumed, while areas far from the refineries experienced no change in SO_2 levels. We exploit this temporal event by estimating difference-in-differences models, using areas far from the refineries as a control group. We focus on the health of newborns as the outcome of interest, both because this is a particularly sensitive group with much policy interest and because birth outcomes are strong predictors of a wide range of future outcomes (Black et al., 2007; Currie, 2009)[2].

While this is not the first pollution-health study to use the closing of an industrial process or other exogenous event as a natural experiment [3], there are several

2. As noted in Joyce & Goldman (1988a) and Chay & Greenstone (2003), focusing on infants also offers a methodological benefit because cumulative exposure can be readily assigned, circumventing issues around mobility and prior exposure.

3. While there are a wide range of studies on this topic using quasi-experimental techniques (see the review in Zivin & Neidell (2013)), the most closely related are Ransom & Iii (1995), Hanna & Oliva (2011), and Currie *et al.* (2013), who all focus on the closing of industrial processes.

Figure 2.1: Adjusted SO_2 levels by proximity to refineries



Notes: SO_2 levels are adjusted by weather variables, the local unemployment rate, and month and year dummy variables. The red lines indicate the approximate dates of the strike. 'Refinery areas' are census tracts where refineries are located, and 'non refinery areas' are census tracts without refineries.

important features of our design that make this an important contribution, mostly centered on parameter identification. First, a common concern in such analyses is that individuals sort into residential locations based, in part, on the amount of air pollution and the employment opportunities in the area, making pollution exposure an endogenous variable [4]. A permanent change in pollution levels can lead to a temporary disequilibrium in the housing market whereby there is no sorting at the time the shock occurs, but sorting is likely to resume as time from the shock passes. If the "post-shock" period includes a long enough time period, then sorting, and hence the endogeneity of pollution, remains a potential concern. In our case, the

4. The link between employment opportunities and pollution endogeneity arises because industry creates both jobs and pollution.

closure of the refineries was a temporary event - lasting approximately one month - making it unlikely that households relocated in search of new employment opportunities or because of preferences for air quality. Second, seemingly exogenous events, such as a strike, may lead to unobserved behavioral changes in the treatment group that affect health, potentially invalidating the research design. Two features make this unlikely in our setting. One, although the variation in pollution is due to the closure of refineries at specific locations, the strike that caused this was a nationwide one centered on pension reforms, with the oil refineries an "unlucky recipient" of the protests. Therefore, any common responses to the strike are accounted for by including a control group. For example, changes in time allocation or activity choice because of the strike affected not only refinery workers but nearly all workers throughout the country [5].

Third, studies that examine the effect of prenatal insults often seek to uncover the distinct effects from different stages of the pregnancy in order to encourage the optimal use of prenatal care. In particular, shocks that occur early in pregnancy, specifically for women who are not yet aware they are pregnant, may leave little opportunity to engage in health-promoting behaviors (Almond & Currie, 2011). In the case of pollution, relatively simple behaviors, such as altering the amount of time spent outside, can yield significant improvements in health (Neidell, 2009)[6].

5. Note that this strategy does not account for avoidance behavior, i.e., changes in time allocation in direct response to the changes in pollution ((Neidell, 2009)). This does not introduce a bias per se but changes the interpretation of estimates, so that our estimates reflect the effect of the strikes net of avoidance behavior. SeeZivin & Neidell (2013) for more details.

6. For example, air quality alerts, which seek to warn the public of dangerous air quality levels, are

Reliably estimating the separate contribution from each trimester is complicated by the fact that pollution levels are often highly correlated across the three trimesters of pregnancy, resulting in severe multicollinearity. Because the strike led to a sharp decrease in pollution for roughly one month, upon which it returned to baseline levels almost immediately after, our research design allows us to overcome this multicollinearity concern to more precisely investigate the separate effects by trimester. Lastly, the handful of quasi-experimental economic studies examining the impact of emissions from energy sources typically focus on the consumption of energy (Currie & Walker, 2011);(Beatty & Shimshack, 2011); (Moretti & Neidell, 2011); (Schikowski *et al.*, 2008). While this consumption side represents an important externality, the production externality is empirically distinct, but has received limited attention[7]. More reliable estimates of the health impacts from energy production are an important component in the development of policies surrounding energy production (Parry & Small, 2005) and the siting of industrial plants. Using this natural experiment, we first demonstrate that although SO_2 is considerably higher in areas close to the refineries, it falls significantly during the strike compared to areas far from the refineries, with regression results supporting the pattern in Figure 1. We find no evidence of changes in two other pollutants, particulate matter and nitrogen dioxide, around the time of the strike, a finding consistent with the change in SO_2 coming from the oil refineries. Turning to health outcomes, we find that birth weight and gestational age of newborns living in the same census tracts as the refineries

particularly targeted at pregnant women.

7. Furthermore, the common pollutants from energy consumption are carbon monoxide and particulate matter.

increased by over 3 and 1.5 percent, respectively, during the strike. Nearly all of the improvement in weight gain can be attributed to the increase in gestation. Furthermore, these effects are primarily driven by exposure during the third trimester of pregnancy, a time when most fetal weight gain occurs. Overall, our estimates suggest that the effects from oil production that accrue to newborns alone are quite sizeable and should be an important part of policy discussions surrounding the production of energy.

2.2. Background: Refineries, Air pollution and Health

2.2.1. Pollution and the refinery closure

Refineries are responsible for 20 percent of SO_2 release in France (Soleille, 2004). Oil refineries convert crude oil to everyday product like gasoline, kerosene, liquefied petroleum. Crude oil contains relatively high quantity of sulfur, which leads to the creation of sulfur dioxide when crude oil is heated at the refinery to produce fuel. The refining process also releases a large number of chemicals such as benzene, chromium and sulfur acid into the atmosphere, which limits our ability to conduct a proper instrumental variable analysis. France has 11 refineries that produce 89 million tons of petrol every year. The main 4 refining companies operating in France are Total, Shell, Esso and Ineos, located in the regions of Haute Normandie, Provence Alpes Côtes dAzur, Rhône-Alpes, Nord-Pas-de-Calais, Pays-de-la-Loire, Ile de France and Alsace. Total refineries are allowed to emit up to 3,500 tons of sulfur dioxide per year which corresponds to 9.6 tons a day. Due to protests over pension reform, protesters successfully ceased production in October, 2010 by mass picketing and the creation of physical blockades around fuel depots. As a result, production was reduced to a minimum or completely ceased for nearly 18 days until the strike was resolved. Closing a refinery is a complex process that requires anywhere from 2 days to one week according to the size of the refinery, and a comparable time period to re-open. Thus, the reduction in SO_2 is likely strongest between mid October and the beginning of November. We focus on the 4 refineries that completely shut down as a result of the strike [8].

2.2.2. *Pollution and health*

Sulfur dioxide (SO_2) is one of a group of highly reactive gasses known as oxides of sulfur (SOx). The largest sources of SO_2 emissions are from fossil fuel combustion at power plants and other industrial facilities (EPA 2011). SO_2 is a colorless gas with a very strong smell. In France, the threshold for SO_2, fixed by the European Act of 2002-13 related to air quality, is 132 parts per billion (ppb) per hour; violations occur when this standard is exceed more than 24 times a year. In comparison, the Clean Air Act in the United States set the one-hour SO_2 standard at 75 ppb, where a violation occurs if the 99th percentile of 1-hour daily maximum concentrations, averaged over 3 years, exceeds this value. This standard was recently strengthened in June 2010, suggesting the need for reliable estimates of the relationship between

8. These refineries are Donges, Feyzin, Gonfreville l'Orcher and Petite Couronne.

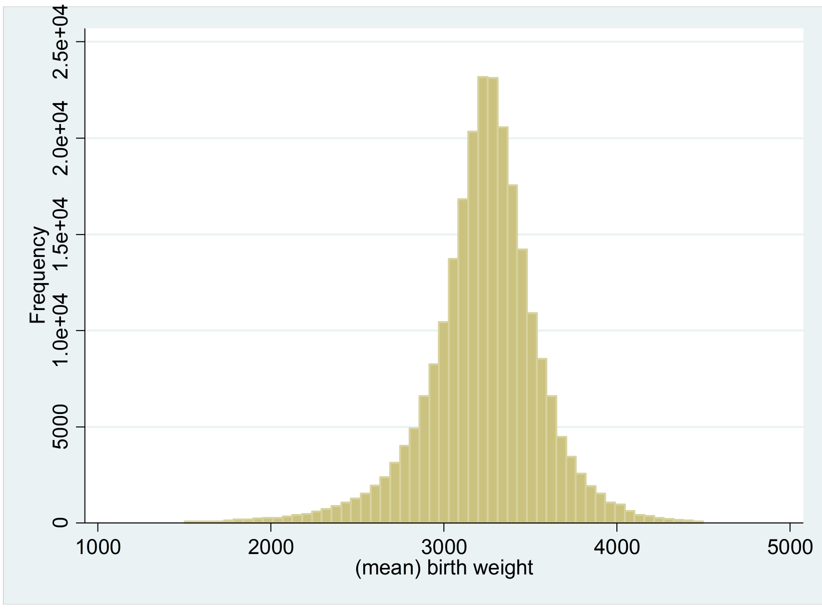
SO_2 and health. Given the rapid stages of development that a fetus goes through in a short period of time, negative shocks can results in both immediate and latent effects (Almond & Currie, 2011). Pollution is one potential shock because it can impair the health of the mother, indirectly compromising fetus health, or cross the placenta, directly affecting the health of the fetus. Slama et al. (2008) describe more extensively possible biological mechanisms by which air pollutants may affect birth outcomes: SO_2, in particular, can harm the fetus by impacting blood viscosity and endothelial function. These changes can affect placental blood flow, transplacental oxygen and nutrient transport, all of which may affect fetal health. Furthermore, while there is a growing consensus that prenatal exposure to pollution affects birth outcomes, there is little understanding about the most susceptible periods of prenatal exposure. While the fetus experiences important organ developments in the first trimester, suggesting a particularly vulnerable stage, the fetus also gains the most weight during the third trimester, suggesting another crucial stage. Evidence from the fetal origins hypothesis suggests that exposure to negative shocks during early pregnancy has no effects at birth but latent impacts later in life (Almond & Currie, 2011), while exposure during late pregnancy is more likely to affect birth outcomes (Stein et al. 2003; Schultz, 2010). . Consistent with this, (Deschenes *et al.*, 2009) find that the sensitivity of birth weight to temperature is concentrated almost entirely in the second and third trimesters of the pregnancy. Whether these same patterns hold for pollution is largely unknown. While not focused on SO_2 per se, several economic studies have found robust evidence that prenatal exposure to pollution affects infant health ((Currie *et al.*, 2009); (Sanders & Stoecker, 2011); (Currie & Walker, 2011)). While most of these studies focus on the effect from exposure during the entire pregnancy, an important contribution of our study is the ability to precisely estimate the effects from exposure during each trimester. Furthermore, previous studies typically focus on pollution stemming from vehicular or industrial emissions, such as particulate matter and carbon monoxide, and our focus on oil refining is more relevant for SO_2.

2.3. Data and empirical strategy

2.3.1. Data sources

Health data are drawn from the French National Hospital Discharge Database (PMSI) from 2007 to 2011. The key variables for our analysis are the year and month of birth, the census tract of residence of the patient, and the birth weight and gestational age at birth. Panel A of table 1 shows the birth weight and the gestational age by month, year and census tract. We also consider low birth weight (<2500 grams) and short gestational age (<37 weeks) as two clinically relevant outcomes. We observe from table 1, panel A that the birth weight and gestational age are lower in census tracts with refineries (the treatment group) than in census tracts without refineries (the control group) for all periods of the study, hinting at potential effects from living near a refinery. Figure 2 shows the distribution of birth weight. Unlike the US, there is much less variation in birth weight in France, a finding consistent with universal access to health care.

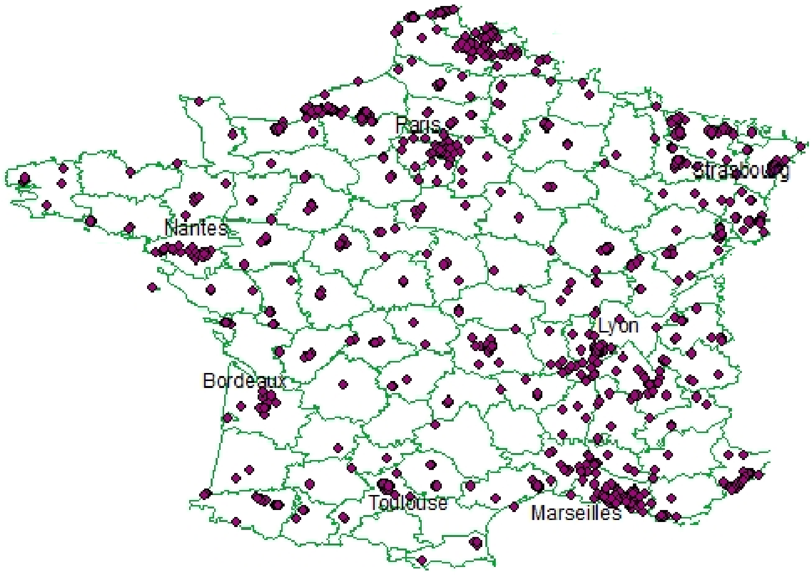
Figure 2.2: Birth weight distribution



Air quality is monitored throughout France by 38 approved air quality monitoring associations (AASQA). The French monitoring station system has approximately 700 measurement monitors equipped with automatic instruments. Figure 3 shows the location of monitoring stations, departmental boundaries (one of the three levels of government below the national level, between the region and the commune), and major cities throughout France.

Not surprisingly, monitors are more highly clustered in major cities. The monitors also show broad coverage of the country, with nearly every department having at least one monitor. We obtain daily measure of ambient air pollution concentrations in microgram per cubic meter (μg/m^3) for all air quality monitors in France for 2007-2010 from the Ministry for Ecology, sustainable development and spatial

Figure 2.3: Air quality monitoring stations and department boundaries in France

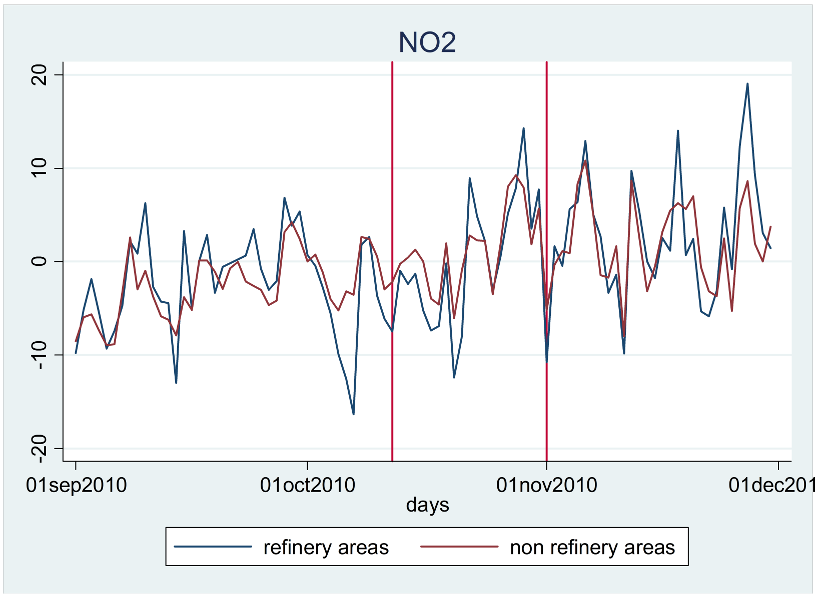


planning (ADEME) database. We also know the exact geographic location of each monitor. Since our main focus is on SO_2, we only include monitors that continuously measured SO_2 during this time period. This leaves us with 187 monitors that span 57 departments and 2864 census tracts. Monthly pollution concentration data are presented in Panel B of Table 1. The most notable aspect of this panel is that SO_2 levels are nearly 4 times higher in areas near the refinery, while the levels are virtually identical for PM10, NO_2 and slightly higher for benzene [9].

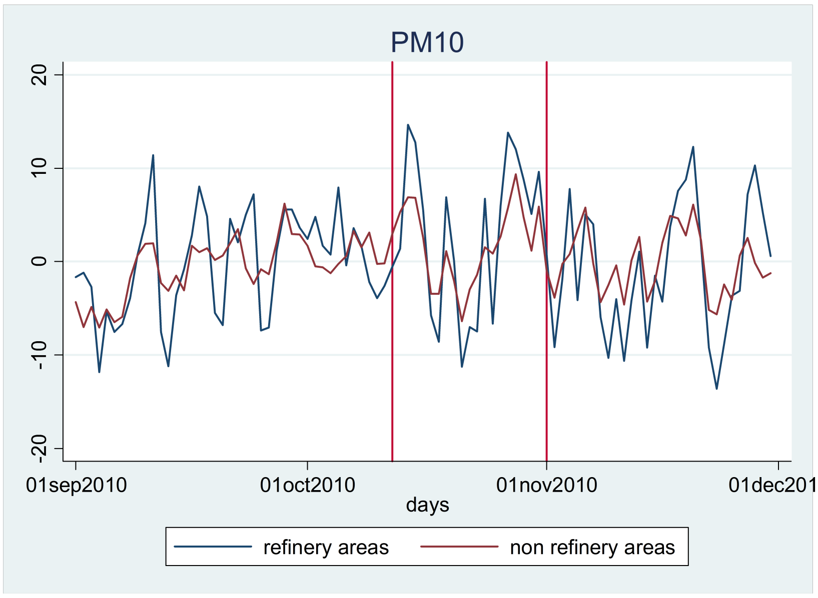
We also present the fraction of days in which the values recorded at the monitors exceeded health standards for SO_2 and PM_{10} [10]. While the number of exceedances is quite low for SO_2 (occurring less than 1% of the time), census tracts with refineries

9. Note that we dropped one inexplicably high measure of benzene (18.44) in order to make the scale of Figure 5 (below) easier to interpret. This measure occurred in a treated census tract on September 25, 2011, so including it would further reinforce the idea that the refineries may affect benzene levels as well.
10. There is no 24 hour air quality standard for benzene and NO_2. Although there is an hourly standard for NO_2, we were only able to obtain daily data.

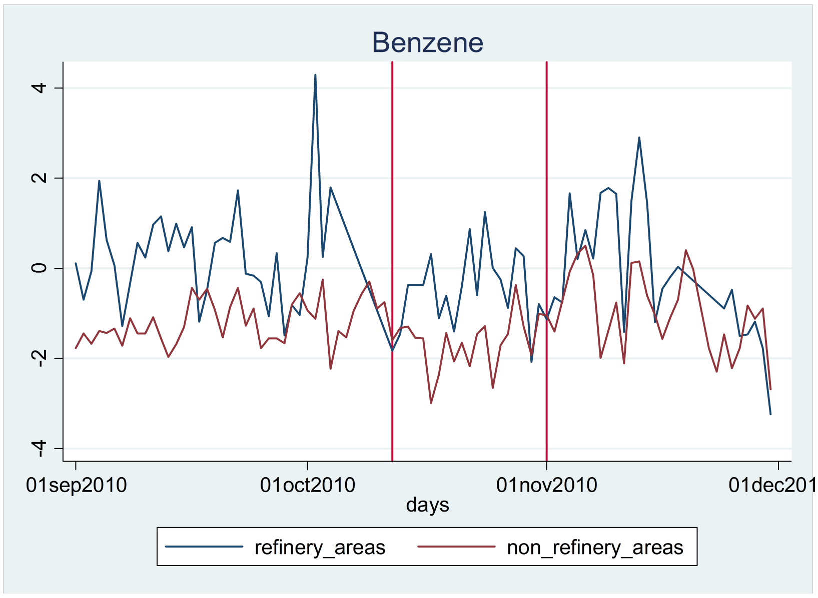
Figure 2.4: Adjusted NO_2, PM10, and Benzene by area



are nearly 10 times more likely to have a violating monitor, consistent with higher SO_2 levels. The rate of exceedances for particulates is much higher on average, occurring nearly 5% of the time, though the rate of violations is quite similar across areas. Since weather has direct effects on health and also affects pollution formation, we also include meteorological data in our analysis. Our weather data come from Meteo France, the French national meteorological service. There are 100 monitors, one in each department. We also have daily measures at each monitor, along with data on the geographic location. We use average and maximum temperature, precipitation, maximum speed wind, prevailing wind direction, and maximum and minimum relative humidity. Summary statistics for daily and monthly measures of weather are presented in Panel C of Table 1. Although we include census tract fixed effects in our regression, which controls for all time invariant characteristics, we also include one measure of economic well-being to capture time varying factors: the



unemployment rate. We use the quarterly rate of unemployment from the National Institute of Statistics and Economic Studies, which is available at the census tract level. Panel D of Table 1 also presents summary statistics for this variable.



Notes: Pollution levels are adjusted by weather variables, the local unemployment rate, and month and year dummy variables. The red lines indicate the approximate dates of the strike.

Table 2.1: Summary statistics

	All	Treatment group (all time periods)	Control group (all time periods)
Outcomes			
birth weight (grams)	3228 [353]	3220 [272]	3228 [354]
birth weight $<$ 2500 grams	.03 [.17]	.02 [.13]	.03 [.17]
gestational age (weeks)	38.86 [1.50]	38.78 [1.31]	38.86 [1.50]
gestational age $<$ 37 weeks	.08 [.27]	.08 [.27]	.08 [.27]
Pollution			
SO_2 - monthly average ($\mu g/m^3$)	3.82 [4.53]	12.87 [10.86]	3.63 [4.10]
SO_2 - % days exceeding AQS	.15 [3.89]	1.284 [1.13]	.12 [.35]
NO_2 - monthly average ($\mu g/m^3$)	24.22 [14.81]	23.35 [12.27]	24.23 [14.83]
PM_{10} - monthly average ($\mu g/m^3$)	22.55 [7.45]	22.52 [6.76]	22.56 [7.45]
PM_{10} - % days exceeding AQS	4.55 [2.09]	5.85 [2.35]	4.47 [2.07]
Benzene - monthly average ($\mu g/m^3$)	1.86 [.80]	2.78 [.74]	1.54 [.25]
Covariates			
mean temperature (řC)	11.85 [5.91]	11.74 [5.65]	11.86 [5.92]
max. temperature(řC)	16.57 [6.76]	16.34 [6.55]	16.57 [6.76]
precipitation (mm)	2.19 [1.44]	2.28 [1.30]	2.19 [1.44]
wind speed (m/sec)	6.94 [1.45]	6.96 [.97]	6.95 [1.45]
wind direction (wind rose)	208.4 [40.4]	205.8 [38.1]	208.4 [40.4]
min. humidity (%)	55.07 [12.03]	56.32[12.50]	55.07 [12.03]
max. humidity (%)	92.7 [4.13]	93.19 [4.02]	92.6 [4.13]
unemployment rate (%)	8.71 [2.22]	8.50 [1.10]	8.71 [2.22]

a

a. Notes: Reported values are means with standard deviations in brackets. The number of observations is 151,624. Air quality standard (AQS) for SO2 is 0.04 ppm (105 $\mu g/m^3$) for every 24 hour period and for PM10 is 50 $\mu g/m^3$ for every 24 hour period.

2.3.2. Merging data

Using the exact location of pollution and meteorology monitors and the census tract of residence for the birth outcomes, we assign pollution to census tracts in a two-step procedure. When a census tract has a pollution monitor in it, we assign that pollution concentration to the census tract. When it does not, we assign pollution using an inverse distance weighted average (IDWA) of pollution, similar to Currie & Neidell (2004). To do this, we compute the centroid of each census tract, and then compute the distance from the centroid to each monitor within the department. We then take the weighted average of pollution measurements from all monitors within a certain distance from the census tract centroid, using the inverse of the distance as weights. We vary the cutoff distance to assess the sensitivity of our results to our assignment of pollution. Although we have a daily measure of pollution and meteorology, health outcomes are only observed at a monthly level. We begin by aggregating pollution and meteorology at a monthly level. Since we only know the month of discharge for newborns, and their average length of stay in the hospital is 5.5 days, we must approximate their date of birth, and thus exposure to the strike. We assume all births occurred on the 1st day of the month, and assign pollution and meteorology from the previous 9 months (we also assess the sensitivity of results by assuming the 15th of the month). For example, an infant discharged in November is born anywhere from October 25th to November 25th, and we assume the birth date is November 1. We then assign exposure to this infant as the mean for the months from February through October, breaking it into 3 month intervals for examining trimester effects.

2.3.3. *Empirical Methodology*

Our goal is to assess the impact of oil production on both pollution levels and health outcomes at birth. We estimate difference in difference models to exploit the unexpected shutdown in production as a result of the strike in October 2010, using areas close to the refineries as the treatment group and areas far from the refineries as the control group. We implement this by estimating the following equation:

$$Y_{cm} = \beta * strike_m * close_c + \delta * X_{cm} + \sigma_m + \alpha_c + \epsilon_{cm} \tag{2.1}$$

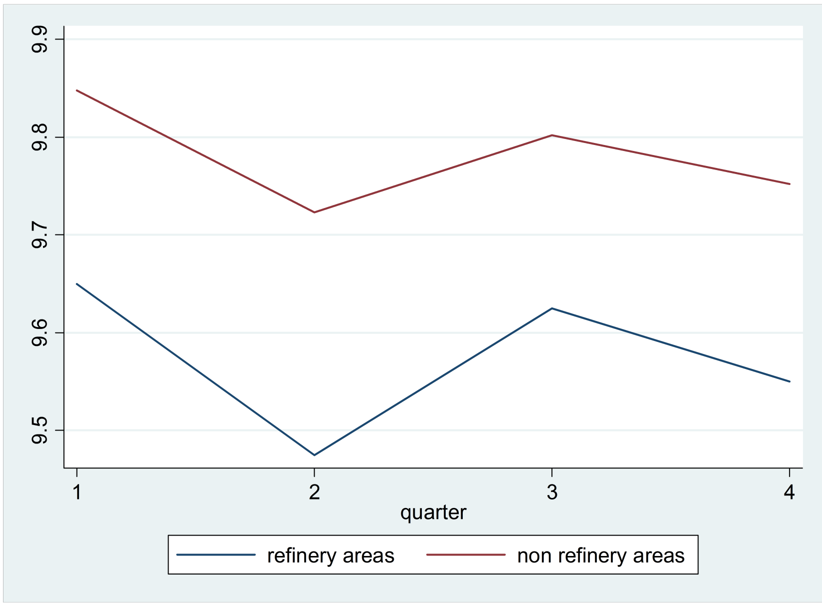
where Y is either ambient pollution concentrations or birth outcomes in census tract c at month m. 'strike' is an indicator variable for the October 2010 period when the strike occurred, and 'close' is an indicator variable for whether the refinery is in the same census tract as the air pollution monitor or patient's residence. β is the difference-in-difference parameter. Xcm is a vector of census tract controls that include weather controls and the quarterly unemployment rate. We control for seasonal and temporal patterns by including month dummies and year dummies in σ_m. We include census tract fixed effects (α_c) to control for time-invariant characteristics of the census tract. ϵ_{cm} represents the error term, which consists of an idiosyncratic component and a term clustered on the department and month. As with any difference in difference design, the key underlying assumption for identification is that the control group serves as a valid counterfactual for the treatment group with parallel trends. Although we can not explicitly verify this assumption, we feel this threat is limited in this setting for several reasons. Because the strike was nationwide, and not just for the workers at oil refineries, any changes in response to the strike likely happened on a global scale that would have affected both the treatment and control groups. Moreover, the strike was a temporary condition, making it unlikely that workers relocated in search of new employment opportunities. Furthermore, because workers in France have health insurance regardless of employment status, there was unlikely to be a change in prenatal care consumption during the time of the strike. Figure 4 provides evidence to support the parallel trends assumption. Since there is little economic data available at such high temporal and spatial resolution, we plot the unemployment rate, which is available quarterly at the census tract, over time for the treatment and control groups. Although the unemployment rate is lower in census blocks with refineries, there is no trend difference between census blocks with refineries and their counterparts, supporting our contention that there are no differential trends across the two groups.

2.4. Results

2.4.1. Refinery closures and pollution levels

We start by examining the effect of strikes on air pollution. The previously mentioned Figure 1 provides a daily graph of adjusted SO2 pollution from September to December, 2010 for the treatment and control groups, with SO2 adjusted by Xcm

Figure 2.5: Unemployment distribution by proximity to refineries



Notes: Unemployment rates are available at the quarterly level for each census tract. 'Refinery areas' are census tracts where refineries are located, and 'non refinery areas' are census tracts without refineries.

and ?m. Prior to the strike, SO2 levels are considerably higher in census tracts with refineries. However, during the strike, SO2 dramatically falls in refinery areas to levels comparable to non-refinery areas. Immediately after the strike, SO2 levels in refinery areas again exceed those of non-refinery areas. This visual display clearly demonstrates a strong, temporal effect of the strike on SO2 levels.

Table 2 provides regression estimates of (1), which are largely analogous to this Figure. In order to gauge the extent of confounding, we successively add more time-varying controls, namely the weather variables and the unemployment rate. Consistent with Figure 1, the strike causes a statistically significant drop in SO2 levels for areas close to refineries. SO2 levels drop during the strike by roughly 15 μg/m^3. Adding controls for weather (column 2) and unemployment (column 3) has no noticeable effect on our estimates.

Table 2.2: The effect of the strike on SO_2 levels

	1	2	3
A. All census tracts			
strike	-15.24*	-15.30*	-15.27*
	(8.796)	(8.799)	(8.772)
Observations	151,624	151,624	151,624
R-squared	0.758	0.758	0.758
B. Census tracts < 8km from monitor			
strike	-16.48*	-17.06*	-16.63*
	(9.020)	(9.065)	(8.713)
Observations	16,945	16,945	16,945
R-squared	0.757	0.758	0.758
C. Census tracts < 2km from monitor			
strike	-26.49**	-28.86**	-25.22**
	(11.23)	(11.30)	(10.79)
Observations	5,652	5,652	5,652
R-squared	0.756	0.757	0.757
weather		x	x
local economic conditions			x

[a]

a. Note: This table provides the coefficient estimates of the effect of strike on Sulfur Dioxide (SO2). All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. The weather variables include average and maximum temperature, precipitation, minimum and maximum humidity, wind speed and direction. The unemployment rate is our measure of local economic conditions. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

The second and third panels explore the effect from different approaches for assigning pollution from monitors to census tracts. Limiting the sample to census tracts within 8 km of a monitor, shown in panel 2, leads to a slight increase in the effect of the strike on SO2 levels. We see a much bigger increase, though still not a statistically significant difference, when we limit to census tracts with 2 km of a monitor. This increase is consistent with a more precise measure of pollution from using a closer monitor. Overall, the results from Table 2, supporting the findings from Figure 1. Figure 5 presents the same plot as Figure 1 for three additional pollutants: NO2, PM10, and benzene. While NO2 and PM10 do not appear to change in response to the strike, Benzene shows a pattern consistent with being affected by the strike, though less stark than that for SO2. While these patterns suggest SO2 is the pollutant most affected by the strike, the possible relationship for other pollutants precludes us from conducting a proper instrumental variable (IV) analysis where we instrument SO2 levels using the strike, though we cautiously provide IV estimates.

2.4.2. *Refinery closures and birth outcomes*

Given that we have found a relationship between the oil refinery strikes and pollution levels, we now turn our attention to the impacts of the strikes on health at birth. Tables 3 and 4 present results of the impact of exposure to the strikes anytime during pregnancy on birth weight and gestation, respectively. The top panel explores the effect on birth weight using the continuous measure and the low birth weight indicator, whereas the bottom focuses on gestational age and short gestation. Within each of the 4 dependent variables, we also explore sensitivity to controls as with the SO2 results, as well as sensitivity to monitor-census tract distance assumptions. For birth weight, we find that birth weight increases by roughly 75 grams during the strike. This result is also insensitive to the addition of weather variables and unemployment. Compared to the mean birth weight of 3228 grams, this represents a 2.3 percent increase in birth weight. If we assume that the only pollutant affected by the refinery is SO2, we can compute the effect of SO2 on birth weight by dividing the effect of the strike on birth weight by the effect of the strike on SO2 as shown in Table 2, akin to instrumental variables (IV). This procedure suggests that a 1

μg/m^3 decrease of SO2 for one month increases birth weight by 5 grams, though we must interpret this with caution because, as noted above, the refineries may have affected other pollutants, such as benzene, which would make IV valid.

Table 2.3: The effect of the strike over the entire pregnancy on birth weight

	1	2	3	4	5	6
		birth weight (g)			birth weight < 2500 g	
A. All census tracts						
strike	73.61*	76.47*	76.44*	-0.020*	-0.021*	-0.021*
	(44.61)	(44.75)	(44.73)	(0.011)	(0.011)	(0.011)
Observations	121,157	121,157	121,157	121,157	121,157	121,157
R-squared	0.053	0.053	0.053	0.066	0.066	0.066
B. Census tracts < 8km from monitor						
strike	71.87	74.87*	74.03*	-0.019*	-0.020*	-0.019*
	(44.50)	(44.91)	(44.83)	(0.011)	(0.012)	(0.012)
Observations	14,169	14,169	14,169	14,169	14,169	14,169
R-squared	0.043	0.044	0.045	0.066	0.067	0.067
C. Census tracts < 2km from monitor						
strike	92.38*	99.43**	99.03**	-0.025*	-0.026*	-0.026*
	(47.21)	(48.41)	(48.41)	(0.014)	(0.014)	(0.014)
Observations	4,962	4,962	4,962	4,962	4,962	4,962
R-squared	0.055	0.059	0.060	0.049	0.054	0.054
weather		x	x		x	x
local economic conditions			x			x

[a]

a. Note: This table provides the coefficient estimates of the effect of exposure to the strike at any time during pregnancy on birth weight. All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

Using an indicator for low birth weight, we find that the strike lowered this rate by roughly 2 percentage points, which is also statistically significant and robust to additional controls. When we limit the distance from pollution monitor to the census tract to 8 km, our estimates change minimally, as with the SO2 results. Limiting to 2 km leads to a larger improvement in birth weight, though the difference is again not statistically significant.

For gestational age, we find similar qualitative results. Using all census tracts, regardless of distance to a pollution monitor, we find the strike increased gestational age by roughly 0.37 weeks, or 2.5 days, which is a 1% change from the baseline

Table 2.4: The effect of the strike over the entire pregnancy on gestation

	1	2	3	4	5	6
		gestation (wks)			gestation < 37 wks	
A. All census tracts						
strike	0.361*	0.382*	0.383*	-0.091***	-0.094***	-0.094***
	(0.194)	(0.196)	(0.195)	(0.029)	(0.029)	(0.029)
Observations	90,134	90,134	90,134	90,134	90,134	90,134
R-squared	0.071	0.071	0.071	0.075	0.075	0.075
B. Census tracts < 8km from monitor						
strike	0.366*	0.373*	0.373*	-0.088***	-0.087***	-0.087***
	(0.196)	(0.197)	(0.197)	(0.030)	(0.030)	(0.030)
Observations	10,761	10,761	10,761	10,761	10,761	10,761
R-squared	0.081	0.083	0.083	0.087	0.089	0.089
C. Census tracts < 2km from monitor						
strike	0.375	0.407*	0.400*	-0.062*	-0.066*	-0.065*
	(0.243)	(0.242)	(0.241)	(0.037)	(0.036)	(0.036)
Observations	3,849	3,849	3,849	3,849	3,849	3,849
R-squared	0.111	0.120	0.120	0.111	0.121	0.121
weather		x	x		x	x
local economic conditions			x			x

[a]

a. Note: This table provides the coefficient estimates of the effect of exposure to the strike at any time during pregnancy on gestation. All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

mean. This yields an IV estimate of a 1 $\mu g/m^3$ decrease of SO2 for one month increases gestational length by 0.18 days. The strike reduces the probability of short gestation by .08. These results are again insensitive to additional controls. While the results do not become larger when limiting to a shorter distance from the census tract to the pollution monitor, the differences are again not statistically significant. To compare the estimates for birth weight and gestation, we perform the following calculation. Since the fetus gains about 200 grams in weight per week in the final month of pregnancy (Cunningham et al., 2010), , the 0.37 week increase in gestation translates into an extra 74 grams in weight, which is nearly identical to our estimate on the impact on birth weight. Therefore, it appears that the reduction in birth weight is solely due to shorter gestation, rather than growth retardation. Since the strike only lasted for less than one month, as previously mentioned one of the advantages of our study is the ability to more precisely isolate the effects by trimester. Table 5 presents results by including exposure to the strike by trimester. We focus solely on census tracts less than 8 km from a monitor and with the meteorological and economic covariates included, though results are robust to different assumptions regarding these choices.

Table 2.5: The effect of the strike on birth weight and gestational age by trimester of pregnancy, census tracts within 8 km of pollution monitor

	1 birth weight (g)	2 birth weight < 2500 g	3 gestation (wks)	4 gestation < 37 wks
strike - 3rd trimester	151.2***	-0.024**	0.847***	-0.110***
	(50.15)	(0.012)	(0.226)	(0.031)
strike - 2nd trimester	10.63	-0.019	0.133	-0.082***
	(66.14)	(0.012)	(0.300)	(0.030)
strike - 1st trimester	60.02	-0.015	0.138	-0.069**
	(78.78)	(0.012)	(0.250)	(0.033)
Observations	14,169	14,169	10,761	10,761
R-squared	0.045	0.067	0.083	0.089
weather		x	x	
local economic conditions			x	

[a]

a. Note: This table provides the coefficient estimates of the effect of strike on birth weight and gestation by trimester of pregnancy when the distance from the census tract to the pollution monitor is less than eight kilometers. All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

We find that almost all of the effects from pollution are due to exposure during the third trimester. Birth weight increases by roughly 150 grams when the strike occurred during the third trimester, which represents a 4.6 percent increase. The effects from the first and, in particular, second trimesters are much smaller and not statistically significant. Turning to the incidence of low birth weight, we find reasonably similar effects across the trimesters, but the third is the largest (and comparable to the estimate for the overall pregnancy) and the only one that is statistically significant. For gestational age, we also find that exposure to the strike in the third trimester has the biggest effect: it increases gestational age by roughly 0.85 weeks, a roughly 2.2 percent increase. This longer gestation translates into roughly 170 grams, which again explains all of the estimated effect on birth weight from third trimester exposure. The effects in the first and second trimester are again much smaller and not statistically significant. Turning to the incidence of short gestation, we again find the third trimester has the biggest effect, but the first and second also appear significantly related to short gestation.

Table 2.6: Estimates using alternative measure of strike exposure, census tracts within 8 km of pollution monitor

	1 birth weight (g)	2 birth weight < 2500 g	3 gestation (wks)	4 gestation < 37 wks
	A. Entire pregnancy			
strike	73.50	-0.0193*	0.347*	-0.0839***
	(44.83)	(0.0114)	(0.198)	(0.0298)
Observations	14,169	14,169	10,769	10,769
R-squared	0.045	0.067	0.083	0.089
	By trimester			
strike - 3rd trimester	148.1***	-0.0224*	0.815***	-0.106***
	(49.92)	(0.0118)	(0.225)	(0.0306)
strike - 2nd trimester	8.323	-0.0184	0.110	-0.0810***
	(66.62)	(0.0121)	(0.307)	(0.0298)
strike - 1st trimester	63.69	-0.0171	0.111	-0.0643*
	(79.75)	(0.0121)	(0.249)	(0.0336)
Observations	14,169	14,169	10,769	10,769
R-squared	0.045	0.067	0.083	0.089

[a]

[a] Note: This table provides the coefficient estimates of the effect of strike on birth weight and gestation assuming all births occurred on the 15th of the month (as opposed to 1st). All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

As previously mentioned, we do not know the exact date of birth of the child, only the month of discharge from the hospital. In Table 6, we present results assuming the date of child's birth is on the 15th of the month instead of the 1st, again focusing solely on the census tracts within 8 km of a monitor. Our results from this specification are virtually identical to the main results, suggesting the lack of knowledge about the exact birth date is not hindering inference. Since pollution and other environmental confounders often show strong seasonal patterns, we want to ensure that our results are not driven by this phenomenon. To assess this, we present estimates from a falsification test where we assign the date of the strike to have occurred on October, 2009, a year before the actual strike occurred. Shown in Table 7, we find that the placebo strike is neither associated with SO2 levels or any of the birth outcome measures. Of the 17 coefficients shown, only 1 is statistically significant (at the 10% level), which is almost exactly what we expect given the chance of a Type I error.

Table 2.7: Effect of placebo strike in October, 2009, census tracts within 8 km of pollution monitor

	1 SO2	2 birth weight (g)	3 birth weight < 2500 g	4 gestation (wks)	5 gestation < 37 wks
A. Entire pregnancy					
strike	0.112	44.02	0.023	-0.149	0.050
	(1.608)	(68.54)	(0.038)	(0.379)	(0.071)
Observations	16,945	14,169	14,169	10,761	10,761
R-squared	0.619	0.045	0.067	0.083	0.089
B. By trimester					
strike - 3rd trimester		69.28	-0.013	0.044	0.040
		(95.03)	(0.009)	(0.350)	(0.115)
strike - 2nd trimester		65.82	-0.018*	0.186	0.049
		(96.34)	(0.010)	(0.437)	(0.112)
strike - 1st trimester		-3.07	0.101	-0.676	0.060
		(151.70)	(0.107)	(0.903)	(0.113)
Observations		14,169	14,169	10,761	10,761
R-squared		0.045	0.068 0.083	0.089	

[a]

a. Note: This table provides the coefficient estimates of the effect of a placebo strike occurring October, 2009 on SO2, birth weight and gestation. All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The distance to the refinery should also be an important factor of the effect on health. In table 9, we modify the treated group by including census tracts within a certain distance to the closest refinery. As shown in the first block, when the treated group corresponds to census tracts within 2 km of the refinery, the positive effect of the strike on birth weight and gestational age is strong and highly significative. Estimates for birth weight are reducing but are not significative anymore when we increase the treated group by including census tracts within 5 km of the refinery. However, gestational age estimates remains significative with a lower coefficient. Increasing the treated group to census tracts within 10 km of the refinery removes significativity. Results are in line with the intuition: the effect on health tends to disappear further away we are from the source of pollution.

Table 2.8: Estimates using alternative distance between census tracts and refineries

	1 birth weight (g)	2 birth weight < 2500 g	3 gestation (wks)	4 gestation < 37 wks
A. treated group: census tracts around 2 km				
strike	87.22***	-0.0223**	0.338**	-0.0636**
	(27.29)	(0.0106)	(0.154)	(0.0248)
Observations	121,157	121,157	90,134	90,134
R-squared	0.044	0.056	0.062	0.065
B. treated group: census tracts around 5 km				
strike	20.22	-0.00842***	0.0879	-0.0240
	(17.05)	(0.00313)	(0.0956)	(0.0315)
Observations	121,157	121,157	90,134	90,134
R-squared	0.044	0.056	0.062	0.065
C. treated group: census tracts around 10 km				
strike	-19.71	0.00471	-0.00286	-0.00709
	(14.91)	(0.00560)	(0.0615)	(0.0158)
Observations	121,157	121,157	90,134	90,134
R-squared	0.044	0.056	0.062	0.065

[a]

a. Note: This table provides the coefficient estimates of the effect of strike on birth weight and gestation using alternative distance for the treated group between refinery and census tract . All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

2.5. Conclusion

The goal of this paper was to examine an externality from energy production, focusing on health impacts as measured by birth outcomes. To account for the en-

Table 2.9: Estimates using alternative measure of exposure to air pollution

	1 SO2	2 birth weight (g)	3 birth weight < 2500 g	4 gestation (wks)	5 gestation < 37 wks
A. Entire pregnancy					
strike	-32.15***	91.75*	-0.0268*	0.365*	-0.0923***
	(49.95)	(0.0146)	(0.195)	(0.0297)	
Observation					
R-squared	0.811	0.063	0.077	0.072	0.076
B. By trimester					
strike - 3rd trimester		156.3***	-0.0274*	0.810***	-0.0996***
		(56.70)	(0.0148)	(0.219)	(0.0301)
strike - 2nd trimester		20.66	-0.0275*	0.0815	-0.0900***
		(67.91)	(0.0146)	(0.307)	(0.0293)
strike - 1st trimester		98.26	-0.0255*	0.204	-0.0871***
		(83.25)	(0.0151)	(0.243)	(0.0313)
Observations					
R-squared					
[a]					

a. Note: This table provides the coefficient estimates of the effect of strike on SO2, birth weight and gestation. All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

dogeneity of pollution exposure, we exploit the oil refinery strike that occurred in October 2010, which led to a sharp, temporary reduction in SO2 in areas close to the refineries. This reduction led to a robust increase in birth weight and gestation of infants, particularly those who were exposed during their third trimester of pregnancy. To gauge the magnitude of these estimates, we perform the following illustrative calculations, similar to Currie *et al.* (2009). We value the improvements in birth weight by computing the percentage change in birth weight from the change in pollution in October, 2010 by dividing the estimated impact of third-trimester SO2 on birth weight from Table 5 (140) by the mean birth weight in our sample (3220) from table 1. We multiply this by the estimated elasticity between birth weight and earnings of 0.1 from Black et al. (2007) to obtain the percentage change in earnings during the month of strikes. We then multiply this by the average gross annual earnings of all full time workers (33,168 euros) from the Directorate for Research, Studies, and Statistics in 2010 in France. Finally, we multiply by the total number of births in 2010 (832,799) to get the change in earnings per year. This gives an estimated increase in nationwide earnings of 120 million euros. Assuming a 40 year working career with 3 percent annual rise in earnings and a 6 percent discount rate, this amounts to 2.933 billion euros per cohort. If we attribute all of the estimated 15 unit decline in SO2 to the strike, this implies that a 1 unit decrease in SO2 increases future earnings of a given birth cohort by 196 million euros per year [11]. While only meant to be illustrative, these estimates suggest that the externalities from oil production that accrue to newborns alone are potentially quite sizeable and should be an important part of policy discussions surrounding the production of energy.

11. Clearly, these estimates understate the full benefits from a decrease in SO2 because they only capture the earnings impacts for a birth cohort and only capture the effects on births. A 3% (4%) discount rate would yield an earnings increase of 328 (272) million euros per 1 unit change in SO2.

Chapter 3

The price of pollution and health: an hedonic approach

This paper examines the impact of a reduction in sulfur dioxide concentration (SO_2) in France on both health outcomes and property prices, at a municipality level, from 2008 to 2011. The paper aims to compare people's willingness to pay for perceived differences in environmental attributes and the real cost in terms of health they are exposed to. To do so, I conduct a hedonic price method analysis using the recent closure affecting one oil refinery in the north of France, in September 2009, as a natural experiment. This contribution shows, first, that a long term shut down in the refining process leads to a reduction in sulfur dioxide concentration. I then use this exogenous shock to assess the impact of a change in air pollution concentration on hospital respiratory admission and property prices. The estimates suggest that the hedonic approach may not always reflect the true environmental health risks.

3.1. Introduction

The hedonic approach to evaluation aims to estimate the economic value of a good using implicit price of the product attributes. In the field of environmental valuation, the economic value attributed to natural resources comes from characteristics of goods which are sold on the market. Rosen (1974) hedonic model shows that the willingness to pay for a change in natural resource can be inferred by the explicit price of a property . In this context, the hedonic approach uses the housing market to estimate the economic benefits of air quality. Chay & Greenstone (2005) find that the elasticity of housing values with respect to particulates concentrations ranges from -.20 to -.35. Greenstone & Gallagher (2008) look at areas chosen for the

Superfund-sponsored cleanups of hazardous waste sites program compared to their counterparts. They find that Superfund cleanups are associated with economically small and statistically insignificant changes in residential property values, property rental rates, housing supply, total population, and types of individuals living near the sites. The estimates from Hanna (2007) suggest that being a mile closer to a polluting manufacturing plant reduces house values by 1.9%.

Besides, the literature focuses also on the relationship between health and housing prices. The superfund cleanups of hazardous waste sites has also been used to shed light on its impact on health status at birth (Currie *et al.*, 2011). Davis (2004) measures the effect of health risk on housing values by exploiting an isolated county in Nevada where residents have experienced a severe increase in pediatric leukemia. The estimated MWTP to avoid pediatric leukemia risk is used to calculate the value of a statistical case of pediatric leukemia. Really recently, Currie *et al.* (2013) look at the housing market and health impacts of 1600 openings and closings of industrial plants that emit toxic pollutant. The paper shows housing values within on mile decrease by 1.5 percent when plants open, and increase by 1.5 percent when plant close.

Whereas, there is a wide literature about the link between pollution and hedonic prices, the literature does not put simultanately light on both the willingness to pay to avoid negative externalities reflected in housing price differential and the real cost in term of health people are exposed to. As Portney (1981) suggested it may be possible to draw inferences about individuals' valuations of risk by combining estimates of the effect of air pollution on both property values and human health risks. This paper aims to analyze to what extend the hedonic price method capture people's willingness to pay for perceived differences in environmental attributes (perceived health risk from pollution and perceived environmental amenities). Individuals' valuations of the lower health risk and any effect on neighborhood aesthetics may be reflected in the price differential associated with proximity to the refinery. The difference between the health benefits and the price differential may be interpreted as a combination of individuals' valuations of neighborhood aesthetics and people perception of a decrease in health risk. and how it may differ from the objective and biological environmental health risk. However, valuations derived from hedonic prices functions must be interpreted carefully. The literature emphasizes a wide number of critics of the Hedonic price analysis.

Under perfect information, the price differential associated with proximity to hazardous sites reflects both individuals' valuations of the greater health risk and any effect on neighborhood aesthetics. Although a biological health risk may exist, it is not clear the extend to which individuals are fully and correctly informed about the health impacts of air pollution. Although individuals may be aware about air pollution, it is less likely they correctly incorporate this risk into their pricing decisions for housing. Imperfect information suggests that the hedonic approach underestimates the true health cost of air pollution whereas damage- function may tend to overestimate the health costs because mortality may be too high (Delucchi *et al.*, 2002). Zabel & Kiel (2000) emphazise it is still unclear how individuals process air quality information when determining their willingness-to-pay for housing.

Another controversial issue is that of market segmentation. Feitelson, Hurd, and Mudge (1996) noted that in theory, hedonic price studies do not require the segmentation of housing markets. However, in practice, several types of market segmentation are likely to exist in most markets. This is because housing markets are not uniform (Adair, Berry, McGreal, 1996; Fletcher, Gallimore, Mangan, 2000). Hence, it is unrealistic to treat the housing market in any geographical location as a single entity. Unfortunately, the definition, composition, and structure of submarkets have not been given much attention in the hedonic-price literature, although it is an important empirical issue.

Another issue frequently associated with the hedonic price model is pollution endogeneity . Hedonic method considers pollution is an exogeneous variable in the regression of housing prices. This is not always correct. Industrial facilities, sources of pollution, are probably located in areas with low population density and relatively low housing prices. On one hand, employees from an industrial company are willing to live close to their place of work so as to limit their everyday transport. On the other hand, atmospheric pollution reduces air quality and the attractivity to live nearby pollution sources. In this context, property prices should decrease.Bajari *et al.* (2012) suggest that ignoring bias from time-varying correlated unobservable considerably understates the benefits of a pollution reduction policy. When differences between locations are imperfectly measured and covary with health risk and housing prices it becomes difficult to disentangle the price effects of health risks from the price effects of other locational amenities(Davis, 2004).

As developped by Chay and Greenstone (2005), differences in terms of pollution preferences may also lead to autoselection bias. In fact, people with low preferences for air quality may sort themselves into location with a high level of air pollution. Preferences for environment are also different across individuals. Households may sort themselves to locations endowed with amenities that match their preferences. The subsample studied may not be representative of the whole population. In this case, hedonic estimation only reflects marginal prices of air quality for a part of the population which do not value air pollution. The value of marginal price of air pollution will be underestimated.

Another issue frequently associated with the hedonic price model is the misspecification of variables. Misspecification is the situation where an irrelevant independent variable is included (over-specification), or where a relevant independent variable (attribute of a product) is omitted (under-specification). As the hedonic price model deals with the implicit prices of quantities of attributes of a product, the problem of misspecification of variables is inevitable. Over-specification gives estimated independent variables that are both unbiased and consistent, but inefficient because of the inclusion of the irrelevant variable, whereas under-specification results in estimated coefficients that are both biased and inconsistent. Measurement errors may also arise if proxy variables are used in the hedonic price model when actual data are unavailable. Consequently, the results generated will be biased and inconsistent. According to Butler (1982), since all estimates of hedonic price models are to some extent misspecified, models that use a small number of key variables generally would suffice.

To mitigate these problems and to infer the impact on housing prices and health status from a reduction in pollution, I use a quasi-experimental approach. The analysis focuses on Dunkirk, a French municipality in the Nord-Pas de Calais region in France where residents have recently experienced a refinery closure. Besides, the conditions of supply and demand are relatively similar in the Nord-Pas de Calais property market such that I can expect a similar set of implicit prices in Dunkirk and in its surrounding analyzing flats and housing separately. Pollution, health outcomes and housing prices are compared before and after the closure with the nearby municipalities with at least one monitoring station, within 50 kilometers, acting as a control group. People living Dunkirk may have a risk behavior which can explain a low health status. The stop in the refining process allows to well measure the effect of pollution on population exposed to pollution coming from the refinery compared to the population living far from the refinery and not exposed to its pollution. The use of variation in pollution and health risk over time is of particular interest to control for unobserved differences across locations. Using a really rich and exhaustive dataset about property transactions, I am able to shed light on population sorting. Moreover, Kuminoff *et al.* (2010) suggest that large gains in accuracy can be realized by moving from the standard linear specifications for the price function to a more flexible framework that uses a combination of spatial fixed effects, quasi-experimental identification, and temporal controls for housing market adjustment. Taking these elements into account, I will analyze in this study the link between the reduction in air pollution, due to the refinery closure, hospital respiratory outcomes and the property value.

This paper aims to be a global study on air pollution, health outcomes and housing prices. I first show that the closure of the refinery leads to a reduction in air pollution; I use this reduction in air pollution to infer the impact it has on hospital respiratory outcomes. In parallel, I analyze to what extend the refinery closure has an impact on property prices to observe the willingness to pay for perceived improvement in air pollution concentration. This valuation may be of use in determining if the willingness to pay for the perceived pollution reduction is in line with the real benefits, the reduction of air pollution has on health.

3.2. Pollution, health and refinery closure

3.2.1. SO_2 pollution and health

This paper focuses on sulfur dioxide (SO_2), one of the major pollutant emitted by oil refineries and the main pollutant from industrial pollution. Sulfur dioxide (SO_2) is one of a group of highly reactive gasses known as oxides of sulfur (SO_x). The largest sources of SO_2 emissions are from fossil fuel combustion at power plants and other industrial facilities (Agency, 2011). SO_2 is a colorless gas with a very strong smell. SO_2 is subject to transformation in the atmosphere and can react with other compounds to form small particles. These particles go deeply to lungs and can cause or deteriorate respiratory diseases, such as emphysema and bronchitis.

Subjects exposed to SO_2 showed decreased lung functioning for children and increased respiratory symptoms for adults (World Health Organization (WHO), 2011), asthma crisis and ocular rash (Pierre Lecoq, 2009). Inflammation of the respiratory tract causes coughing, mucus secretion, aggravation of asthma and chronic bronchitis and makes people more prone to infections of the respiratory tract (World Health Organization (WHO), 2011). The effects seem stronger for high levels of exposure and people with asthma are more sensitive to SO_2. The number of hospital admissions for cardiopathy and mortality increases on days with high SO_2 air concentration (Finkelstein *et al.*, 2003). Human clinical studies consistently demonstrate respiratory morbidity among exercising asthmatics following peak exposures

(5-10 min) to SO_2 concentrations equals 0.4 ppm, with respiratory effects occurring at concentrations as low as 0.2 ppm in some asthmatics (Organization, 2005).

The chemical life of SO_2 is around two days, much lower than the monthly temporal period used in the study. If it was the time of exposure which causes the pathologies and hospitalizations, I could risk identifying a change in the composition of the local population rather then the effect of the refinery stop. In fact, employees of the refinery very exposed to pollution could leave Dunkirk leading to a decrease in the number of admission and the severity of disease. Nevertheless, the impact of SO_2 is relatively local and fast such as the concentration observed in the month m does not depend on the concentration of the previous month, m-1. Thus, the effect of the refinery activity stop is well identified.

3.2.2. Pollution and the refinery closure

Refineries are responsible for 20 percent of SO2 release in France (Soleille, 2004). Oil refineries convert crude oil to everyday product like gasoline, kerosene, liquefied petroleum. Crude oil and coal contain relatively high quantity of sulfur. SO_2 is creating when crude oil or coal is heated at the refinery to produce fuel. Thus, the refining process releases a large number of chemicals such as benzene, chromium and sulfur acid into the atmosphere. in this context, refineries are considered upper tier SEVESO sites for most of their activities. In Europe, the 'Seveso' directive applies to around 10,000 industrial establishments where dangerous substances are used or stored in large quantities, mainly in the chemicals, petrochemicals, storage, and metal refining sectors. The Seveso Directive obliges Member States to ensure that operators have a policy in place to prevent major accidents.

The Flandres refinery, close to Dunkirk in northern France, is part of SEVESO sites. The refinery produces liquefied gases (propane and butane), fuel for airplanes and automobiles (gasoline and diesel), domestic and industrial fuel and biofuels. Its refining capacity is up to 7.8 million tons/annum. The refinery employs nearby 370 employees on average, which represents 0,4% of the global employment in Dunkirk. The refinery also annually works with 775 establishments which accounts for 87 million euros, among which 275 establishments are localized in Nord-Pas-de-Calais, which accounts for 44,1 million euros. Most of the employees live near their workplace: two thirds of the employees of the refinery live in Dunkirk. Most of them works full-time, under a permanent employment contract and are labor workers. One quarter of the refinery workers belongs to intermediate profession (INSEE, 2010). In september 2009, the production of the refinery has been shutted down due to poor demand and margins. Given the poor outlook, french oil giant Total announces definitely the closure of its refinery in 2010. I reasonably believe that the closure had an impact on all the Nord-Pas de Calais region affecting not only refinery workers but also a large range of subcontractors from all the Nord-Pas de Calais region. On the one hand, not surprisingly, figure 1 shows that Dunkirk where the activity of the refinery shutted down experienced a decrease in sulfur dioxide (SO_2) concentration after 2009. On the other hand, census districts without refineries experienced minimal pollution changes as the variation seems to stay stable overtime.

3.2.3. The incidence of the refinery closure on housing prices

The hedonic price model, derived mostly from Lancaster (1966) consumer theory and Rosen's (1974) model, shows that a differentiated good can be described by a vector of its characteristics. In the case of a property, these characteristics may include structural attributes (e.g., number of bedrooms), neighborhood public services (e.g., local school quality), and local environmental amenities (e.g., presence of a toxic site). People have the opportunity to select the combination of features they prefer, given their income. In this context, areas with elevated health risks such as Dunkirk must have lower housing prices to attract potential homeowners. As emphasized by Greenstone & Gallagher (2008), the price differential associated with proximity to hazardous sites reflects both individuals' valuations of the greater health risk and any effect on neighborhood aesthetics.

The case of the refinery closure in Dunkirk is compounded of two effects: economics and environmental. First, a refinery closure represents an economic shock in term of employment and economic activity in the surrounding of Dunkirk. This economic shock has affected the entire Nord-Pas de Calais region in terms of economic outcomes due to direct and indirect employment of the refinery.

Second, the environmental shock is represented by a decrease in pollution concentration in the area of Dunkirk following the refinery closure. Thus, I assume that the closure of the refinery may reduce health risks from air pollution, likely increase the aesthetic value of proximity to the site and finally, enhance the value of neighbourhood properties.

While the hedonic price analysis focuses on the perception people have about a reduction in air pollution, the damage function is an objective measure of the impact of air pollution on health. The hedonic price analysis will only capture people's willingness to pay for perceived differences in environmental attributes, and their direct consequences. And the individual's perception of health risk may differ from the real risk population living near toxic sites is exposed to. From this criteria, the article aims to analyze diagnosis differences between both methods. In this context, it may be attractive to directly compare the willingness to pay to avoid negative externalities reflected in housing price differential and the real cost in term of health people are exposed to. However, on one hand, the damage function is likely to underestimate the total health benefit because it omits effects on mortality. A large number of studies of the economic value of reducing SO2 and related pollutant find a high part of the quantified health benefits associated with mortality reduction. More than 80% of monetized benefits were attributed to reductions in premature mortality (Krupnick *et al.*, 2002). Health benefits are orders of magnitude larger than the ones we can report in this study. In addition, disutility from health is only one part of people perception. Not only the health risk due to air pollution may be a source of disutility but also environmental amenities: the view of a factory chimney, the smoke or possible odor may have an impact on the perception of the population living near a refinery. On the other hand, the hedonic prices analysis ignores the effect on the entire property market and the long term effect. While there are a number of constraints that prevent any comparison between both methods, I rather put in lights estimation results of environmental benefits from both methods in the following sections.

3.3. Dataset presentation

3.3.1. Pollution data

Air quality is monitored throughout France (mainland and overseas departments) by 38 approved air quality monitoring associations (AASQA). The French monitoring station system counts approximately 700 measurement stations equipped with automatic instruments and nearly 400 experts implement this monitoring system. I focus on sulfur dioxide (SO_2) concentration in the Nord-Pas de Calais region around 50 kilometers from Dunkirk which represents two departments, geographical level below the regional level. I obtain daily measure of ambient air pollution concentrations in microgram per cubic meter ($\mu g/m^3$) for all air quality monitors in France for 2008-2011 from the Ministry for Ecology, sustainable development and spatial planning (ADEME) database and more recently from the national institute of industrial environment and risks (INERIS). Sulfur dioxide concentration after the refinery closure in Dunkirk has decreased from 12.65 to 6.58 $\mu g/m^3$ whereas SO_2 concentration in the other municipalities, has decreased from 3.12 to 2.21 $\mu g/m^3$ in average. As a consequence, the difference in difference between Dunkirk and the control group after the refinery closure is 5.16 $\mu g/m^3$.

Table 1 presents the summary statistics of all the variables. Monthly pollution concentration data are presented in panel A of the summary statistics where I present a measure of expected exposure to SO_2 after having removed stations that do not exist for the entire period: from 2008 to 2011. In addition I remove monitoring stations that do not measure SO_2. Note also that only some municipality dispose of a monitor. These 2 departments represent 238 municipalities and 16 air pollution monitoring stations[1]. The distribution of monitoring stations throughout France is represented in figure 4 with a marquee which represents the area of the study, the Nord-Pas de Calais. The summary statistics indicates that the monthly SO_2 concentration decreases after the refinery closure. The level is quite low over the period due to the monthly aggregation from daily data. The SO_2 concentration after the refinery closure decreases from 12.65 to 6.58 $\mu g/m^3$ whereas the SO_2 concentration in the others municipalities reduces from 3.12 to 2.21 $\mu g/m^3$. As a consequence, the difference in difference between Dunkirk and its control group after the refinery closure is 5.16 $\mu g/m^3$.

1. There is still a difference in observations because I face some missing data.

Table 3.1: statistics summary Mean[SE]

Variables	The entire period	Before sept. 2009	After sept. 2009
	Panel A : Pollution concentration N=185,687		
Sulfur dioxyde(SO_2) ($\mu g/m^3$) mean average	2.120748[1.787319]	2.478793 [1.995779]	1.485385[1.078837]
	Panel B : Health outcomes		
Number of respiratory admission by inseecode, month, age and sex	.0586082 [.2804725]	.0568436 [.276021]	.060765 [.2858043]
Length of stay(days)	6.140569 [8.302577]	6.173942 [8.536914]	6.102398 [8.026011]
	Panel C : Weather variables		
Precipitations (mm)	2.008765 [1.015537]	2.087354 [.8599679]	1.869306 [1.232342]
Max_Temp (^0C)	14.59093 [6.513577]	14.25712 [6.044898]	15.18327 [7.233453]
Av_Temp (^0C)	10.73392 [5.627273]	10.4839 [5.27849]	11.17759 [6.173196]
wind_speed (m/sec)	7.827492 [1.123382]	7.956915 [1.17471]	7.597826 [.9849877]
Wind_direct (rose des vents)	206.6803 [44.91593]	211.3447 [40.79121]	198.4032 [50.37515]
Min_Humidity (%)	60.00654 [11.78418]	60.55584 [10.71448]	59.0318 [13.41971]
Max_Humidity(%)	93.76189 [2.770849]	92.83191 [2.65916]	95.41219 [2.120282]
	Panel D : socio-economic variables		
Age (in days)	13389.47 [12241.85]	13389.47 [12241.86]	13389.47 [12241.86]
Unemployment (%)	12.2314 [2.432976]	11.5939 [2.346794]	13.01057 [2.3063]
	The entire period	Before 2010	After 2010
	Panel E : Flat variables N=2848		
price_ttc	128914.1 [63255.87]	126400.6 [59737.58]	131217.9 [66253.21]
level_number	1.958224 [1.745731]	1.800306 [1.711344]	2.111111 [1.765608]
typ_flat	.7847612 [.4110601]	.8193833 [.3848416]	.7530283 [.4313953]
terrace	.1264045 [.332363]	.1666667 [.3728149]	.089502 [.2855631]
attic	.0551264 [.2282669]	.0594714 [.2365917]	.051144 [.2203656]
balcon	.2247191 [.4174705]	.2129222 [.4095234]	.2355316 [.4244735]
parking	.6646106 [3.454972]	.9608856 [4.093888]	.3405973 [2.543684]
garden	.0400281 [.1960595]	.041116 [.1986314]	.039031 [.1937338]
house_srf (m^2)	61.05585 [27.84797]	61.59224 [26.06801]	60.57463 [29.35488]
less_5_years	.2977528 [.4573505]	.3230543 [.467815]	.2745626 [.4464438]
room_nb	2.594101 [1.308262]	2.709251 [1.20337]	2.48856 [1.389549]
	Panel F : House variables N=13870		
price_ttc	152684.1 [97988.77]	150242.5 [68782.41]	154983.5 [119076.9]
pool	.0021629 [.0464588]	.0017839 [.0422012]	.0025199 [.0501393]
typ_house	.3647441 [.4813757]	.4148952 [.4927406]	.3175136[.4655417]
terrace	.0487383 [.2153281]	.0404341[.1969897]	.0565589 [.2310139]
attic	.1312906 [.33773]	.132154 [.3386833]	.1304774 [.3368515]
balcon	.0023071 [.047979]	.0019325[.0439211]	.0026599 [.0515096]
parking	.2772987 [3.591735]	.2660619[3.460214]	.2879751 [3.712596]
house_srf (m^2)	107.198 [42.3946]	106.4439[40.30485]	107.9031[44.25248]
less_5_years	.0515501 [.2211249]	.0567861 [.2314506]	.0466191 [.2108363]
room_nb	3.888745 [2.429112]	3.998365 [2.3414]	3.785494[2.504704]

[a]

a. Note: This table indicates the mean and standard error for the estimation key variables from 2008 to 2011 in France, before shutting down the refining process and after the shutting down in the north of France.

3.3.2. *Morbidity data*

Health data are drawn from the French National Hospital Discharge Database (PMSI) from 2008 to 2011. The key variables for the analysis are the month of admission , the length of stay and the place of residence of the patient. I dispose of an anonymous summary which gives information about the geographical code of residence of a patient, its age, its sex, its main and linked diagnosis. Pathologies are classified with respect to the international disease classification. I dispose of both outpatient discharge admission and emergency admission. People who did not stay overnight in the hospital have a length of stay of zero in the dataset. I do not dispose of the exact day of admission but I have the length of stay. I will use this information to construct a measure of expected exposure to air pollution. Panel B of the summary statistics sheds light on the number of admissions for respiratory disease by month, year, municipality, age in days and sex to keep the most disaggregated dataset. Note that most of the dataset consist of patients that have stayed less than 15 days (90 percent) at the hospital.

3.3.3. *Weather data and socioeconomic data*

I use temperature, precipitation, humidity and wind data in the analysis to both control for the direct effects of weather on health (Chay & Greenstone, 2003) and also to leverage the quasi-experimental features of wind direction and wind speed in distributing pollution from refineries (Hanna & Oliva, 2011), (Beatty & Shimshack, 2011). The weather data come from Meteo France, the French national meteorological service. I dispose of the average and maximum temperature in Celsius degree, the number of precipitation in millimeters, the maximum speed wind in meters per second, the prevailing wind direction in wind rose and the maximum and minimum relative humidity in percent [2]. I use data a daily frequency from the French weather monitoring system. I dispose nearby one station by department and I mainly use it as a control in the regression. Weather data are presented in Panel C of the summary statistics.

Temperature and the intensity of sunlight play an important influence in the chemical reactions that occur in the atmosphere to form photochemical smog from other pollutants. Besides, wind speed and direction measurements are important for air quality monitoring. If high pollutant concentrations are measured at a monitoring station, the wind data recorded at the station can be used to determine the general direction and area of the emissions. Wind speed can greatly affect the pollutant concentration in a local area. The higher the wind speed, the lower the pollutant concentration. Wind dilutes pollutants and rapidly disperses them throughout the immediate area. Humidity and precipitation can also act on pollutants in the air to create more dangerous secondary pollutants, such as the substances responsible for acid rain. On the opposite, precipitation have a beneficial effect by washing pollutant particles from the air and helping to minimize particulate matter formed by activities such as construction and some industrial processes (Agency, 2012a).

I also use the quarterly rate of unemployment from the National Institute of

2. The relative humidity of an air-water mixture is defined as the ratio of the partial pressure of water vapor in the mixture to the saturated vapor pressure of water at a prescribed temperature.

Statistics and Economic Studies (INSEE).

3.3.4. Property Prices

I use a unique and really rich dataset coming from the chambre des notaires, PERVAL dataset. I dispose of every property transactions around 50 kilometers from Dunkirk from 2008 to 2011. The control variables that are used in the estimation have really few missing value in order to keep the most exhaustive dataset. Panel E and F of Table 1 presents the main characteristics of the 2848 and 13870 property transactions for flats and houses respectively. The key variables are the property prices, the number of floors, the number of rooms, the type of flat or house, the property surface, the presence of a terrace, an attic, a parking, a balcony, a pool or a garden and a variable which indicates if the property has less than 5 years. I observe from the summary statistics that property prices for both houses and flats increase on average after the refinery closure.

3.4. Estimation

I am looking at the causal relationship between a closure in the refining activity, local pollution levels, the number of contemporaneous respiratory hospitalizations, and the property prices in the North of France.

The purpose, first, is to estimate the impact of the refining closure $post\_closure_{cm}$ on pollution concentration or health outcomes Y_{cm} or the parameter β_1 in the following linear probability model:

$$Y_{cm} = \beta_0 + \beta_1 post\_closure_{cm} + \beta_2 post_{cm} + \beta_3 closure_{cm} + X_{cm} + \theta_m + \omega_y + \epsilon_{cm} \quad (3.1)$$

where the dependent variable Y_{cm} represents either SO_2 pollution concentration or health outcomes within each municipality c at month m according to the model I present. $post_{cm}$ represents the timing after the refinery closure in september 2009. I also include in all regression the variable $closure_{cm}$ coded as 1 if it is Dunkirk to control for time-invariant unobserved covariates of respiratory admissions. The control group coded as zero represents the rest of municipalities in the north of France with at least one monitoring station. $post\_closure_{cm}$ is the difference in difference estimator and represents Dunkirk area after the refinery closure in 2009. Because I dispose of monthly hospital admissions, I aggregate the daily measure of pollution concentration at a monthly level [3] . X_{cm} is a vector of municipality controls that include weather controls W_{cm}. I also control for temporal variation in pollution including month fixed effects θ_m, year fixed effects ω_y to limit the influence of pollution outliers. ϵ_{cm} represents the error term.

This difference in difference model satisfies the following equation:

$$\mathbb{E}[\epsilon_{cm}, post\_closure_{cm}] = 0 \quad (3.2)$$

Similarly, the second objective is to estimate the impact of the refining closure

3. When aggregating daily data, the monthly average concentration take into account monitors without missing data

$post\_closure_{pcy}$ on housing prices P_{pcy} or the parameter α_1 in the following linear probability model:

$$P_{pcy} = \alpha_0 + \alpha_1 post\_closure_{pcy} + \alpha_2 post_{pcy} + \alpha_3 closure_{pcy} + W_{pcy} + \varphi_y + \nu_c + \sigma_{cy} \quad (3.3)$$

where the dependent variable P_{pcy} represents the log price of each property p, within each municipality c at year y. One of the fundamental assumptions of the hedonic price method, as underlined previously, is that households have perfect information. If people are not fully informed of the linkages between the environmental attribute and benefits to them or their property, the value will not be reflected in home prices. Moreover, the hedonic price schedule does not adjust instantaneously to changes in demand or supply conditions in the housing market. Many factors like imperfect information and transaction costs will then result in the process of adjustment taking some time. That is why, I assume a buying decision being made over a year in this model such that I observe the change in housing prices from 2010, three months after the refinery activity stop and ten months before the real closure. I reasonably assume that 2010 takes into account the delay for adjustment to changes in demand and the anticipation of the closure people can have before the refinery definitely shuts down. $post\_closure_{pcy}$ represents Dunkirk after the refinery closure in 2009. W_{pcy} is a vector of property and municipalities controls that include property characteristics and the level of unemployment in each area. I also control for temporal variation in pollution including year fixed effects η_y to limit the influence of pollution outliers. I also include in all regression a municipality fixed effect ν_c to control for time-invariant unobserved covariates of prices. In fact, much of the variation may be explained by unobserved factors that characterize particular properties like geographical features, neighborhood characteristics and design amenities (Davis, 2004). σ_{cy} represents the error term.

The issue of sorting is crucial in this study. Individuals that choose to live near these sites may have a low willingness to pay to avoid the associated health risks. If consumers value the closure of the refinery, then the closure should cause individuals to sort such that there is an increase in the number of people, who place a high value on environmental quality, living near the refinery. Hanna & Oliva (2011)'s paper discuss that the closure may have altered the attractiveness of surrounding neighborhoods for individuals with strong preferences for air quality. If wealthier or healthier people moved closer to the refinery after it closed, the estimates may simply be capturing the differences in labor supply between the old and new residents of the refinery neighborhood. Although focusing on the years around the closure reduces the probability of selective migration, I will analyze the change in population after the closure.

3.5. Results

3.5.1. Pollution concentration and refinery closure

I start by examining the effect of refinery closure on air pollution. The result is interesting in itself to understand to what extend the refining activity influences the amount of pollution released in the air. Figure 1 provides a monthly graph of SO_2 residual estimation in $\mu g/m^3$ from 2008 to 2011 for Dunkirk and municipalities far from Dunkirk. I restrict the sample to the Nord-Pas de Calais region.

I control for seasonal patterns adding year and month dummies to deal with the falling pollution trend and the high variation in air pollution I observe overtime. After the closure of the refinery in Dunkirk, pollution in SO_2 falls in Dunkirk relative to their counterparts where pollution seems to be stable. After the closure, the level of pollution concentration in Dunkirk catches up the level of pollution observed in the other municipalities.

Table 2 details this effect more carefully. I present the estimate of β_1 from Equation (1), where I replace the $post\_closure$ variable by a dummy variable whether the municipality is Dunkirk after september 2009. I consider a simple measure of SO_2 in Micrograms per cubic meter in column 1. While column 2 adds weather control, column 3 repeats the estimation with municipalities fixed effect. Column 4 adds the unemployment variable as a proxy for the economic activity trend. In the last column, I take full advantage of the variation in distance between the municipality and the monitoring station by reducing the dataset to two kilometers distance between the centroid of the municipality and monitors in order to have a more precise measure of exposure to air pollution. Note that all specifications include month, year and municipality fixed effect and are clustered by municipality and month.

Dunkirk, municipality where is the refinery, shows a reduction in SO_2 air pollution after the refinery closure of 5 micrograms per cubic meter in average. The

Table 3.2: first stage regression

VARIABLES	(1) SO2	(2) SO2	(3) SO2	(4) SO2	(5) SO2
post_closure	-5.166**	-5.163**	-4.951**	-4.993**	-5.072**
	(2.146)	(2.150)	(2.166)	(2.188)	(2.214)
post	0.856	1.059			
	(0.627)	(0.643)			
closure	9.547***	9.536***			
	(2.065)	(2.060)			
av_temp		-0.764	-0.358	-0.316	-0.221
		(0.431)	(0.439)	(0.447)	(0.462)
pp		-0.192	-0.179	-0.199	-0.243
		(0.156)	(0.156)	(0.163)	(0.164)
max_temp		0.734	0.365	0.348	0.283
		(0.417)	(0.415)	(0.420)	(0.432)
speed_wind		0.362**	0.321*	0.339**	0.405**
		(0.129)	(0.149)	(0.152)	(0.158)
direct_wind		-0.00301	-0.00306	-0.00324	-0.00395
		(0.00473)	(0.00518)	(0.00521)	(0.00540)
min_humidity		0.00528	-0.0242	-0.0209	-0.0252
		(0.0411)	(0.0414)	(0.0409)	(0.0383)
max_humidity		0.0214	0.0306	0.0461	0.0654
		(0.0886)	(0.0948)	(0.102)	(0.101)
Un				-0.327	-0.317
				(0.251)	(0.253)
Year FE	x	x	x	x	x
Month FE	x	x	x	x	x
municipalities FE			x	x	x
Distance < 2km					x
Observations	185,687	185,687	185,687	185,687	175,212
R-squared	0.478	0.486	0.503	0.504	0.504

[a]

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the concentration of SO2. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

refinery closure substantially reduces pollution and it is consistent with all measure of pollution. The estimate is not driven by standard demographic characteristics (column 3) nor by neighborhood specific trends (municipality FE). Taking distance into account increases the magnitude of the effect and the estimate remains significant at the five percent level.

3.5.2. Respiratory outcomes and refinery closure

I now focus on the health impact evaluation of the refinery closure. Results from the reduced form are important because refinery closure may involve benefits (Hanna & Oliva, 2011) and may help to establish the extent to which refinery closure reduces economic costs in the form of earning profits and wages. Figure 2 and Figure 3 provide a monthly graph of the evolution of the number of respiratory admission and the evolution of the length of stay at the hospital during the period respectively.

By disentangling the extensive from the intensive margin, the study shows to what extend SO_2 air pollution triggers a disease or increase its severity. It is quite clear from the graph that the length of stay in Dunkirk falls after the closure of the refinery in such way that it closes the gap between Dunkirk and the rest of the municipalities. Concerning the number of admissions, I do not observe any obvious changes from the graph. Let's note the number of respiratory admissions is higher in winter due to the relationship between temperature and respiratory admissions. I present the estimate of β_1 from Equation (1) in Table 3, 4 and 5. Note that table 3 uses a logit model as I code the outcome as 1 if there is at least one hospital respiratory admission. The outcome is coded as zero otherwise.

Table 3.3: Respiratory admissions: reduced form regressions

VARIABLES	(1) Admissions	(2) Admissions	(3) Admissions	(4) Admissions
post_closure	0.0216	0.0219	-0.00822	-0.0113
	(0.0396)	(0.0395)	(0.0389)	(0.0398)
post	0.0989***	0.0767***	0.182***	0.184***
	(0.0198)	(0.0277)	(0.0390)	(0.0396)
closure	0.00615	0.00571	-0.0289	-0.0227
	(0.0302)	(0.0301)	(0.0311)	(0.0312)
Un			-0.0883***	-0.0930***
			(0.00503)	(0.00453)
av_temp		0.00362	-0.0166	-0.00798
		(0.0478)	(0.0433)	(0.0412)
pp		-0.00505	-0.00790	-0.00523
		(0.0152)	(0.0167)	(0.0180)
max_temp		0.00662	0.0274	0.0199
		(0.0423)	(0.0402)	(0.0386)
speed_wind		0.00887	0.0139	0.00931
		(0.0134)	(0.0164)	(0.0187)
direct_wind		-1.92e-05	-0.000136	-0.000116
		(0.000314)	(0.000305)	(0.000320)
min_humidity		0.00243	0.00548	0.00453
		(0.00495)	(0.00566)	(0.00592)
max_humidity		-0.00279	0.00179	0.00427
		(0.00885)	(0.0114)	(0.0115)
Year FE	x	x	x	x
Month FE	x	x	x	x
Distance < 2km				x
Observations	185,687	185,687	185,687	175,212
[a]				

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on respiratory admissions. All regressions are estimated using logit, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Adults could adjust their behavior after the closure of the refinery. If they remain unemployed after the refinery closure, they may have more time to go to the hospital for a visit. The length of stay may be a better outcome to control for any avoidance behavior as they cannot have an influence on the number of days they will stay at the hospital once they are admitted. While I cannot rule out the possibility of avoidance behavior on the parts of workers because of reductions in hours worked, I do not find decreases in the number of admissions for respiratory outcomes in table 3.

Nevertheless, I find a slowdown in the severity of illness in table 4 suggesting avoidance behavior is unlikely to drive our results. Length of stay is reduced in average by one day after the refinery closure. The decrease in concentration of air pollution may not be sufficient to reduce the number of respiratory disease whereas it may be enough to have an effect on the severity of a respiratory disease [4].

As previously indicated, children and the elderly are particularly sensitive to pollution. To assess this, I repeat the same regressions stratifying by age group in table 5.

This assessment also helps to probe into potential avoidance behavior. Consistent with this, I only find statistically significant estimates for the youngest (under age 5) and oldest (>70) age group. These differences by age also help to allay concerns regarding avoidance behavior. While workers may have dramatically changed their exposure because they were no longer working, it is unlikely that children changed their behavior. And since nearly the entire elderly group is not working, it is unlikely their behavior changed either. Hence, these results suggest that for at least some groups I am picking up effects net of avoidance behavior. Thus, the estimates suggest

4. The estimation of the length of stay may present a selection bias. In fact, it only considers people already admitted at the hospital. However, the non significativity I observe in the estimation of the admission shows that there is no extensive margin effect

Table 3.4: The length of stay: reduced form regressions

VARIABLES	(1) LOS	(2) LOS	(3) LOS	(4) LOS	(5) LOS
post_closure	-1.083***	-1.082***	-1.087***	-1.096***	-1.104***
	(0.257)	(0.267)	(0.244)	(0.246)	(0.234)
post	-0.138	-0.448			
	(0.336)	(0.355)			
closure	1.062***	1.055***			
	(0.209)	(0.221)			
av_temp		0.269	0.169	0.192	0.184
		(0.327)	(0.328)	(0.326)	(0.362)
pp		-0.255**	-0.240**	-0.249**	-0.281**
		(0.111)	(0.0986)	(0.0981)	(0.111)
max_temp		-0.267	-0.180	-0.191	-0.198
		(0.306)	(0.313)	(0.308)	(0.337)
speed_wind		0.213	0.188	0.194	0.253*
		(0.138)	(0.139)	(0.138)	(0.127)
direct_wind		-0.00116	-0.000673	-0.000725	-0.00184
		(0.00239)	(0.00260)	(0.00261)	(0.00277)
min_humidity		-0.0187	-0.0181	-0.0167	-0.0127
		(0.0364)	(0.0397)	(0.0393)	(0.0420)
max_humidity		0.0127	0.0195	0.0261	0.00988
		(0.0421)	(0.0442)	(0.0451)	(0.0496)
Un				-0.146	-0.139
				(0.200)	(0.202)
Year FE	x	x	x	x	x
Month FE	x	x	x	x	x
municipalities FE			x	x	x
Distance < 2km					x
Observations	18,544	18,544	18,544	18,544	17,514
R^2	0.003	0.003	0.005	0.005	0.006
a					

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the number of respiratory admissions. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Table 3.5: LOS Reduced form regressions by age category

Variables	(1) 0-5	(2) 5-15	(3) 15-25	(4) 25-40	(5) 40-60	(6) 60-70	(7) >70
post_closure	-0.260*	-0.203	0.536	-1.631*	-0.986	-0.851	-2.127***
	(0.122)	(0.578)	(0.683)	(0.744)	(0.759)	(1.048)	(0.490)
Un	0.160	0.0389	-0.159	-0.273	0.389	-1.755***	0.222
	(0.0904)	(0.196)	(0.529)	(0.290)	(0.227)	(0.478)	(0.269)
av_temp	0.148	0.218	-0.104	1.206**	0.418	1.477*	0.515
	(0.177)	(0.321)	(0.311)	(0.432)	(0.872)	(0.738)	(0.510)
pp	-0.119	-0.0678	-0.0497	-0.277	-0.185	-0.572***	-0.464*
	(0.0701)	(0.0826)	(0.181)	(0.200)	(0.235)	(0.175)	(0.220)
max_temp	-0.164	-0.203	0.0139	-1.104**	-0.325	-1.461*	-0.519
	(0.165)	(0.291)	(0.325)	(0.471)	(0.774)	(0.686)	(0.493)
speed_wind	0.0529	-0.0388	0.110	-0.0885	0.0313	0.571**	0.0938
	(0.0872)	(0.109)	(0.176)	(0.274)	(0.251)	(0.236)	(0.175)
direct_wind	0.00153	-0.00174	0.00151	0.00302	-0.00162	-0.00699	-0.00239
	(0.00220)	(0.00312)	(0.00354)	(0.00452)	(0.00321)	(0.00633)	(0.00363)
min_humidity	-0.0148	-0.0189	0.0375	-0.0704	0.0260	-0.121*	-0.0408
	(0.0186)	(0.0481)	(0.0520)	(0.0818)	(0.0619)	(0.0593)	(0.0591)
max_humidity	-0.00741	0.0535	-0.155	0.129	-0.00800	0.146	0.0700
	(0.0312)	(0.0944)	(0.0887)	(0.146)	(0.0835)	(0.102)	(0.0916)
Year FE	x	x	x	x	x	x	x
Month FE	x	x	x	x	x	x	x
Municipalities FE	x	x	x	x	x	x	x
Observations	9,205	1,958	1,882	3,035	6,023	3,375	9,163
R^2	0.024	0.045	0.025	0.025	0.008	0.035	0.016 [a]

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the number of respiratory admissions by age category. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Demographic controls include an indicator for gender, age and unemployment. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

that daily variation in SO2 air pollution has economically significant effects on the severity of respiratory outcomes for at risks population. The stop in activity also reduces the length of stay at the hospital for people between 25 and 40 years old. This last category is part of the working population. I am in line with the literature saying that pollution significantly reduces worker productivity through health. Thus, cleaning up pollution from refining activity would benefit the economy in term of labor. This result suggests that workers with young children are even more likely to miss a day of work because pollution impacts their own health and their children's ones.

3.5.3. Property prices and refinery closure

Flats and houses are presented in two separated models due to their specificity and differences in price evolution. Buying a flat is not the same investment decision than buying an house, and the reasons behind such an investment may differ. I have previously underlined a causal effect between a reduction in pollution concentration and a reduction in the severity of illness. I now wonder how much people are willing to pay for this reduction in SO2 concentration and the improvement in health status I observe. This reduction in pollution and health risk after the refinery closure may also have some effect on property values, reflect of perception, that I look at in this section.

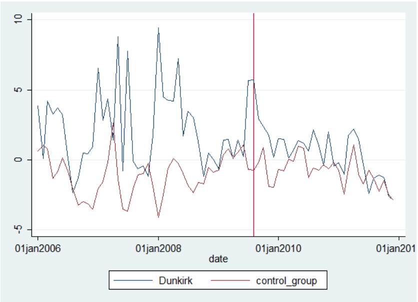
3.5.3.1. House Prices

In this model, the exclusion restriction may be violated if other events happen at the same time of the closure. Nevertheless, it is unlikely to have happen. As far as I know, no other economic events or policies coincided at the exact same time. However the area of Dunkirk is exposed to many economic activity downturns. The economic activity evolves constantly in Dunkirk mainly due to its port. France's third-ranking port, Dunkirk is well known as a port handling heavy bulk cargoes for its numerous industrial installations. Dunkirk represents also the French's leading port for ore and coal imports; the French's leading port for containerized fruit imports; the French's leading port for copper imports; and the French's second-ranking port for trade with Great Britain. The port's territory covers 7,000 hectares and includes ten towns : Dunkirk, Saint-Pol-sur-Mer, Fort-Mardyck, Grande-Synthe, Mardyck, Loon-Plage, Gravelines, Craywick, Saint-Georges-sur-l'Aa and Bourbourg. Thus, I control every estimation for unemploymnent, as a control variable for the economic activity. Furthemore table 10 shows that the refinery closure does not have any impact on unemployment. After the closure, jobs have indeed been offered to the workers of the refinery in other group facilities or unit. Table 10 also introduce *buyer_migration*, *single* and *male* outcome variables which are coded as one if buyers migrate outside Dunkirk, are single and male respectively. The refinery closure does not have any impact on the composition of the population as I do not find any significativity for those coefficients. The closure did not influence buyers to migrate to Dunkirk.

As Currie *et al.* (2013), since in our empirical application, all residents near or far from the refinery live in the Nord-Pas de Calais region, I assume that the wage effects are similar for both nearby residents and those a little further from a plant. In this context, holding all other factors fixed, all of the change in property prices following the shock in Dunkirk compared to its counterparts indicate to what extend individuals evaluate health risk and any effect on neighborhood aesthetics.
Figure 4 shows the evolution of housing prices before and after the definitive refinery closure which occurs in 2010.

The price trend in Dunkirk follows a similar increasing trend before and after the closure whereas the price trend in the others municipalities keeps rising but at a decreasing rate after the closure. Table 6 presents the effect of the refinery closure on the housing prices in more details.

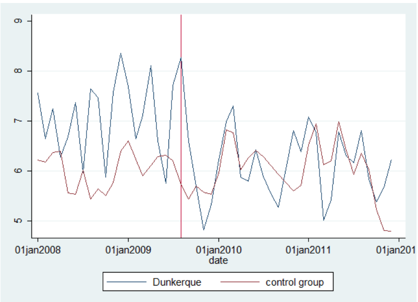
Figure 3.1: Monthly residual regression for SO_2 in ($\mu g/m^3$) from 2006 to 2011



a

a. Notes: This graph represents the SO_2 residual concentration for municipalities with a refinery versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure

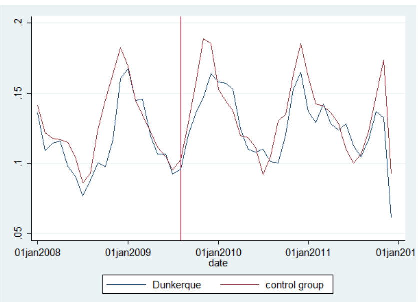
Figure 3.2: Evolution of the lengh of stay in days from 2008 to 2011



a

a. Notes: This graph represents the evolution of the length of stay for the area of Dunkirk versus municipalitiess without a refinery within the same department. September 2009 corresponds to the refinery closure.

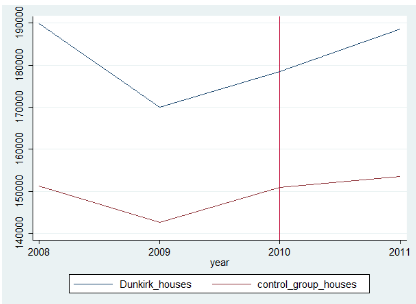
Figure 3.3: Evolution of the number of admissions from 2008 to 2011



a

a. Notes: This graph represents the evolution of the number of admissions for the area of Dunkirk versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure.

Figure 3.4: Evolution of the housing prices from 2008 to 2011



a

a. Notes: This graph represents the evolution of the number of housing prices for the area of Dunkirk versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure.

Table 3.6: Reduced form regressions: HOUSE

VARIABLES	(1) OLS	(2) OLS	(3) FE	(4) FE	(5) distance<10km OLS	(6) OLS	(7) FE	(8) FE
post_treatment	-0.0236	0.0492	0.0556***	0.0505***	0.0144	0.0894**	0.0915**	0.0613**
	(0.0620)	(0.0525)	(0.00517)	(0.00773)	(0.0272)	(0.0239)	(0.0232)	(0.0163)
post	0.0698	0.0451			-0.0335*	0.000681		
	(0.0712)	(0.0663)			(0.0113)	(0.00949)		
treatment	0.220***	0.158***			0.161***	0.114***		
	(0.0469)	(0.0389)			(0.0272)	(0.0169)		
pool		0.283***	0.185**	0.168**		0.570***	0.370**	0.278*
		(0.0567)	(0.0603)	(0.0574)		(0.0844)	(0.0978)	(0.0887)
typ_pav		0.209***	0.188***	0.188***		0.174***	0.154***	0.147***
		(0.0146)	(0.0135)	(0.0151)		(0.0158)	(0.0122)	(0.00476)
terrace		0.113***	0.132***	0.131***		0.146	0.123	0.131
		(0.0147)	(0.0115)	(0.0111)		(0.0900)	(0.0840)	(0.0794)
attic		-0.0420**	-0.00213	-0.00499		0.0327	0.0376	0.00522
		(0.0166)	(0.00683)	(0.00583)		(0.0247)	(0.0526)	(0.0371)
balcon		0.140*	0.0691	0.0672		0.0296**	0.0436	0.0550
		(0.0705)	(0.0416)	(0.0456)		(0.00658)	(0.0278)	(0.0254)
srf_parking		-0.000521	-0.000512	-0.000694		-0.00431**	-0.00260	-0.00415*
		(0.000831)	(0.000728)	(0.000725)		(0.000746)	(0.00165)	(0.00159)
srf_hab_init		0.00508***	0.00487***	0.00532***		0.00364*	0.00349**	0.00531***
		(0.000451)	(0.000429)	(0.000129)		(0.00126)	(0.00105)	(0.000188)
moins5ans		0.257***	0.237***	0.236***		0.244***	0.237***	0.253***
		(0.0173)	(0.0165)	(0.0157)		(0.0295)	(0.0226)	(0.0222)
nbr_pieces		0.0331***	0.0334***	0.0282***		0.0427**	0.0394***	0.0252**
		(0.00522)	(0.00534)	(0.00434)		(0.00963)	(0.00656)	(0.00632)
Un				-0.0332				-0.0373
				(0.0201)				(0.0570)
Year FE	x	x	x	x	x	x	x	x
municipality FE			x	x			x	x
distance<10km					x	x	x	x
Observations	13,870	8,745	8,745	8,499	3,066	1,509	1,509	1,263
R-squared	0.018	0.408	0.510	0.522	0.046	0.380	0.451	0.485
[a]								

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of house. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

While the first block of table 6 looks at the entire dataset, the second block reduces the sample to municipalities located within a distance of 10 kilometers around the refinery to have a more homogeneous dataset in terms of trends (Currie *et al.*, 2013). Column 1 uses a difference in difference model without property characteristics. Column 2 adds property characteristics and column 3 controls for municipality specificity using municipality fixed effect. The last column takes full advantage of the unemployment variable in each municipality to control for any differences in the activity trend between the treatment and control group. Table 6 shows a rise in the price of houses after the closure and it is robust in all fixed effect models even when I present a more homogeneous dataset in the second block. Adding property characteristics and municipality fixed effect improve the specification of the model. Unemployment, the economic activity proxy added in the last column reduces the effect of the shock on houses prices. The 5 micrograms per cubic meter reduction in SO_2 pollution I observe after the refinery closure leads to nearby 5% increase in housing prices.

In view of these estimations, housing buyers seems willing to pay for an improvement in air quality.

3.5.3.2. Flats Prices

The exclusion restriction may still be violated in this model if the refinery closure impacted property prices independently of the health and pollution channel. People living near the refinery may actually be concerned differently by the refinery closure.

In fact, the closure of a refinery not only represents a pollution shock but also an economic activity shock. While population becomes unemployed, the economic dynamism of the area falls. This activity slowdown decreases the wealth of people living in the area where the refinery was. Economic activity can have a direct impact on prices and may bias the estimate. A basic finding in the literature is that income tends to influence willingness to pay for environment positively and significantly (Hokby & Soderqvist, 2003). Willingness to pay for environment may differ with the level of income. In this context, People buying expensive dwellings may be more sensitive to pollution than people buying cheap dwellings.

Table 3.7: Reduced form regressions: FLAT

VARIABLES	(1) OLS	(2) OLS	(3) FE	(4) FE	(5) distance<10km OLS	(6) OLS	(7) FE	(8) FE
post_treatment	-0.120	-0.102**	-0.0771*	-0.0784*	0.0270	-0.0394	-0.0209	-0.0329
	(0.0804)	(0.0304)	(0.0357)	(0.0344)	(0.125)	(0.0822)	(0.0778)	(0.113)
post	0.0417	0.146***			-0.140	0.100		
	(0.0969)	(0.0288)			(0.102)	(0.0619)		
treatment	0.0392	0.0813**			0.0642	0.307**		
	(0.0562)	(0.0234)			(0.0787)	(0.0767)		
nbr_niveau		0.0208***	0.0192***	0.0197***		0.00416	0.00504	0.00587
		(0.00570)	(0.00446)	(0.00460)		(0.00351)	(0.00418)	(0.00434)
typ_as		0.0888	0.0876*	0.0857*		0.147**	0.149**	0.144**
		(0.0586)	(0.0430)	(0.0419)		(0.0381)	(0.0363)	(0.0380)
terrace		0.0931**	0.0346	0.0369		0.0481	0.0503	0.0482
		(0.0351)	(0.0276)	(0.0265)		(0.0345)	(0.0386)	(0.0362)
attic		-0.0599	-0.0538	-0.0519		-0.0552	-0.0597	-0.0562
		(0.0421)	(0.0363)	(0.0359)		(0.0545)	(0.0570)	(0.0559)
balcon		0.146***	0.0691***	0.0671***		0.0348	0.0292	0.0237
		(0.0175)	(0.0183)	(0.0186)		(0.0298)	(0.0322)	(0.0347)
srf_parking		-6.44e-05	-0.000397	0.000141		-0.00162	-0.00204	-0.00207
		(0.00153)	(0.00184)	(0.00210)		(0.00322)	(0.00346)	(0.00338)
garden		0.0578	0.0384	0.0426		0.0775	0.0700	0.0826
		(0.0489)	(0.0489)	(0.0506)		(0.0370)	(0.0356)	(0.0437)
srf_hab_init		0.00838***	0.00993***	0.00991***		0.0103***	0.0104***	0.0104***
		(0.000462)	(0.000469)	(0.000516)		(0.000622)	(0.000602)	(0.000703)
moins5ans		0.485***	0.605***	0.605***		0.529***	0.527***	0.527***
		(0.0815)	(0.0986)	(0.0965)		(0.0680)	(0.0672)	(0.0657)
nbr_pieces		0.0211	0.0223	0.0221		0.0246	0.0254	0.0264
		(0.0129)	(0.0146)	(0.0153)		(0.0147)	(0.0150)	(0.0157)
Un				-0.0704				-0.105**
				(0.0463)				(0.0226)
Year FE	x	x	x	x	x	x	x	x
municipality FE			x	x			x	x
distance<10km					x	x	x	x
Observations	2,848	2,065	2,065	2,026	1,269	875	875	836
R-squared	0.006	0.515	0.630	0.629	0.011	0.620	0.625	0.621
a								

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of flat. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7 and 8 focus on the effect of the refinery closure on the price of flats. Models are similar to the one presented for houses. While table 7 looks at the entire dataset, table 8 split the sample with respect to flat prices. While environmental quality increases, Table 7 shows a reduction in the price of flat after the closure. The result is still robust when I add unemployment as a control in the last column. Note that the significativity is quite low. However, the significativity disappears in the second part of table 7 when I consider a more homogeneous sample with municipalities not further than 10 kilometers away.

Table 3.8: Reduced form regressions: FLAT with respect to prices

VARIABLES	price<mean OLS	(2) FE	(3) FE		price>mean OLS	(6) FE	(7) FE		price>150000 OLS	(10) FE	(11) FE
post_closure	-0.00391	-0.0222	-0.0284*		0.0369	0.0517*	0.0515*		0.0649	0.0661***	0.0647***
	(0.0344)	(0.0189)	(0.0143)		(0.0332)	(0.0253)	(0.0239)		(0.0352)	(0.0156)	(0.0159)
post	-0.0102				0.0637				0.0613		
	(0.0217)				(0.0391)				(0.0464)		
post_closure	0.0692**				0.00327				0.00863		
	(0.0215)				(0.0175)				(0.0277)		
floor_number	0.00933*	0.0127**	0.0133***		0.0150***	0.0134**	0.0134**		0.0204***	0.0195***	0.0193***
	(0.00447)	(0.00383)	(0.00377)		(0.00372)	(0.00480)	(0.00503)		(0.00437)	(0.00424)	(0.00439)
flat_type	0.181***	0.182***	0.182***		-0.0515	-0.0671	-0.0671		-0.0409*	-0.0587*	-0.0583*
	(0.0206)	(0.0327)	(0.0333)		(0.0398)	(0.0398)	(0.0398)		(0.0198)	(0.0254)	(0.0255)
terrace	0.0729	0.0283	0.0323		0.0445*	0.0178	0.0178		0.0509	0.0339	0.0337
	(0.0494)	(0.0470)	(0.0466)		(0.0217)	(0.0156)	(0.0156)		(0.0270)	(0.0224)	(0.0225)
attic	0.00543	0.0112	0.0157		-0.177***	-0.168***	-0.168***		-0.186***	-0.188***	-0.190***
	(0.0302)	(0.0276)	(0.0271)		(0.0227)	(0.0235)	(0.0236)		(0.0284)	(0.0246)	(0.0259)
balcony	0.112**	0.0594**	0.0568**		0.0411**	0.0191	0.0191		0.0319	0.00303	0.00258
	(0.0338)	(0.0227)	(0.0226)		(0.0147)	(0.0136)	(0.0137)		(0.0219)	(0.0187)	(0.0191)
parking_srf	0.00688**	0.00260	0.00290		-0.00362*	-0.00338	-0.00339		-0.00379	-0.00424	-0.00435
	(0.00207)	(0.00263)	(0.00251)		(0.00189)	(0.00270)	(0.00278)		(0.00341)	(0.00380)	(0.00384)
garden	0.0275	0.0328	0.0357		0.0723*	0.0268	0.0267		0.0277	-0.00931	-0.0108
	(0.0462)	(0.0382)	(0.0392)		(0.0350)	(0.0450)	(0.0453)		(0.0333)	(0.0403)	(0.0401)
house_srf	0.00609***	0.00713***	0.00696***		0.00379***	0.00522***	0.00522***		0.00277***	0.00399***	0.00400***
	(0.000584)	(0.000497)	(0.000565)		(0.000432)	(0.000465)	(0.000462)		(0.000431)	(0.000469)	(0.000473)
less_5_years	0.318***	0.332***	0.330***		0.119***	0.218***	0.218***		0.0614**	0.140***	0.139***
	(0.0359)	(0.0708)	(0.0704)		(0.0247)	(0.0230)	(0.0231)		(0.0211)	(0.0261)	(0.0263)
room_nb	0.0231*	0.0301**	0.0311**		-0.00580	-0.00423	-0.00422		0.00623	0.00732	0.00732
	(0.0111)	(0.0119)	(0.0130)		(0.00801)	(0.00903)	(0.00906)		(0.00919)	(0.00942)	(0.00944)
Un			-0.0830**				0.00124				0.0173
			(0.0264)				(0.0263)				(0.0299)
Year FE	x	x	x	x	x	x	x	x	x	x	x
municipality FE		x	x			x	x			x	x
Observations	1,199	1,199	1,161		866	866	865		597	597	597
R-squared	0.417	0.514	0.515		0.248	0.371	0.370		0.263	0.371	0.371
a											

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of flat by level of flat prices. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

In addition, I test for the heterogeneity of the treatment results by splitting the sample into different prices level. In table 8, prices are set below and above the mean. The difference in difference estimator becomes positive and significant in Table 8 when I consider a subsample above the mean. After the closure, for a price of flat set above the mean, people are willing to pay more for the improvment in air pollution. Thus, the level of prices plays a crucial role. This intuition is confirmed if I consider higher prices, set above 150 000 euros: both the significativity of the difference in difference estimator and the size of the impact of the refinery closure on flat prices are increasing. On the contrary, people buying cheaper flats, set below the mean in the subsample of table 8, are less sensitive to pollution than they are to the economic activity which can explain the negative and significative coefficient. I find suggestive evidence that population living in expensive flats have a significant and larger response to pollution than population living in cheaper flats.

Table 3.9: the distribution of flat with respect to social class

price_mean_dummy	executive	intermediate professions	employees	craftsman	factory workers	retired	farmers	others	Total
0	144	320	228	47	124	135	12	15	1,025
	38.71	63.49	74.51	42.73	82.67	37.29	46.15 41.67	54.93	
1	228	184	78 63	26	227	14	21	841	
	61.29	36.51	25.49	57.27	17.33	62.71	53.85	58.33	45.07
Total	372	504	306	110	150	362	26	36	1,866
	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
a									

a. Notes: This table presents the distribution of flat prices with respect to the socioeconomic category

To reinforce these results, table 9 shows the distribution of flats prices by social categories. Most of executives buy expensive flats whereas most of factory workers buy cheaper flats. Factory workers represent the main social category working in oil refinery.

Table 10 tests more carefully for the change in population and demographics following the closure of the refinery. The housing data contains a lot of information concerning the buyer and the seller characteristics so that I am able to observe precisely the change in population due to environmental preferences following the closure. The refinery closure has no effect on migration and buyers characteristics in general (table 10). Migration represents a dummy coded 0 if the last property of the buyer is in the same municipality than the the property he is buying.

Table 3.10: Other effect for the overall sample

VARIABLES	(1) Migration_buyer	(2) Un	(3) Male	(4) Single
post_closure	-0.0197	0.0317	-0.00468	-0.000505
	(0.0277)	(0.125)	(0.0182)	(0.0315)
Year FE	x	x	x	x
municipality FE	x	x	x	x
Observations	16,718	16,159	16,718	16,718
R-squared	0.113	0.964	0.169	0.054
[a]				

a. Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on demographics. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

However, table 11 details the effect of the closure on each social class category. Thanks to a really precise dataset about every transaction, I am able to present results about the impact of the closure on sellers (block 1) and buyers (block 2) separately by social category. Table 11 indicates that after the refinery closure, factory workers among others are significative sellers of their properties in Dunkirk compared to the control group (column 5). On the contrary, executives sell less in Dunkirk after the closure than the control group (column 1). While executive, sensitive to the reduction in pollution prefer to stay in Dunkirk, factory workers, sensitive to the economic shock, prefer to sell their property in Dunkirk. The Intermediate profession group is not significative in both sellers and buyers estimations. Looking at the buyers estimation, the coefficient for factory workers is not significative. The coefficient for executive is negative suggesting that executive are buying less in Dunkirk after the shock than they are in the other municipalities. It suggests executive may decide not to sell their property in Dunkirk following the closure but the reduction in pollution may not be strong enough to be an incentive for executive to come to live in the area. Behind such complexity in population behavior, the monetary evaluation with the hedonic price analysis, alone, cannot give an absolute value (Maslianskaia Pautrel, 2009). Hedonic results can be used by policy makers provided a careful interpretation and/or comparing with others method such as the evaluation of sanitary costs.

Table 3.11: The effect of the closure on social class

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SELLERS								
Variables	executive	intermediate professions	employees	craftsman	factory workers	retired	farmers	others
post_closure	-0.0241***	0.00978	0.0309**	-0.00209	0.0228***	0.0377**	0.00418**	0.0174**
	(0.00286)	(0.00949)	(0.0102)	(0.00477)	(0.00635)	(0.0136)	(0.00130)	(0.00535)
Observations	16,718	16,718	16,718	16,718	16,718	16,718	16,718	16,718
R-squared	0.026	0.032	0.026	0.034	0.046	0.060	0.036	0.034
BUYERS								
Variables	executive	intermediate professions	employees	craftsman	factory workers	retired	farmers	others
post_treatment	-0.0313**	0.0223	0.0344***	0.00860**	0.0167	-0.00498	-0.00624**	0.00524
	(0.0126)	(0.0134)	(0.00915)	(0.00293)	(0.0122)	(0.0113)	(0.00216)	(0.00410)
Observations	16,718	16,718	16,718	16,718	16,718	16,718	16,718	16,718
R-squared	0.034	0.043	0.031	0.036	0.065	0.033	0.040	0.021
a								

a. Notes: This table presents the effect of the closure on the migration by social category. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

3.5.4. Monetary evaluation

I derive an approximation of the cost of pollution in term of labor by looking at the cost of hospital admission [5]. The ExternE project gives a monetary evaluation that I take into account to derive an approximation of the cost of pollution. The cost of one hospital admission and one emergency room visit for respiratory illness is evaluated at 2 000 Euros per admission and 670 Euros per visit respectively (price year 2000) (Bickel et al.,2005). The effect size that I find suggests that $5\mu g/m^3$ decrease in sulfur dioxide pollution I observe after the closure (Table 2) leads to a decrease in the length of stay at the hospital close to one day (Table 4). As one day of hospital respiratory admission costs 2 000 Euros and the yearly average number of hospital respiratory admission in Dunkirk is 700, there is a yearly cost difference of nearby 1 400 000 Euros (2000*700) between before and after the refinery closure in Dunkirk. Note that I underevaluate these costs due to the monthly average I used for our estimation. Individuals' valuations of the lower health risk and any effect on neighborhood aesthetics may be reflected in the price differential associated with proximity to the refinery. The difference between the health benefits and the price differential may be interpreted as a combination of individuals' valuations of neighborhood aesthetics and people perception of a decrease in health risk.

In comparison, the housing prices increase by 5 percent after the refinery closure

5. The ExternE project (External Costs of Energy), project of the European Commission, aims to measure the damages to society which are not paid for by its main actors and to translate these damages into a monetary value.

which suggests an average benefit of nearby 7500 Euros (150242.5*0.05) for each transaction. As the number of housing transactions in Dunkirk is 337 in 2010, the housing market in Dunkirk benefits by more than 2 Million in 2010.

Puting in lights both methods shows first a lack of significative results when the hedonic approach is applied on the overall property market. On the contrary, houses prices, increasing after the refinery closure, are in line with the environmental economic intuition, whereas flats prices, decreasing after the refinery closure, do not comply with the theoretical literature. The increase/decrease of property prices comes from an increase/decrease of the net demand [6] Some elements of explanation may be drawn from those results. Unless I assume flat owners are less informed about pollution impact, objectivity from the damage function versus perception from the hedonic price analysis do not explain results differences between houses and flats. The perception of the decrease in health risk may indeed be smaller than the real decrease in health risk in case of imperfect information. Besides, because their budget constraint is tight, buyers of cheap flats spending constraint may be stronger than others. Even if they were aware about air pollution reduction, environmental effect is larger in relative terms for owners/buyers of cheap flats than the economic effect. On the contrary, environmental effect is larger in relative terms for owners/buyers of houses and expensive flats than the economic effect.

6. This is due to an increase/decrease of the household demand who want to live in the area or an increase/decrease of the houshold demand who want to leave the area

3.6. Conclusion

This paper tests the short term effect of sulfur dioxide (SO2) on respiratory outcomes and the global impact it has on housing prices. Our first goal is to assess the impact of air pollution reduction on health outcomes for those municipalities that experienced a reduction in air pollution following the activity stop of the oil refinery in Dunkirk, north of France, in september 2009. Since I have a panel dataset, the best way to isolate causal effect of the reduction in SO2 concentration from the closure at one oil refinery is to examine outcomes differences between Dunkirk from its counterparts overtime. I look at the effects of a closure on local measures of sulfur dioxide (SO2) concentration. I address several longstanding issues dealing with non-random selection and behavioral responses to air pollution that may bias previous studies. This result is particularly significative for at risks population such as children below 5 years old and people over 70 years old. I also find a significative effect for adults between 25 and 40 years old suggesting air pollution concentration can have a deleterious impact on labor outcomes.

The second part of the project aims to analyze to what extend the willingness to pay for an improvement in air quality reflects the positive effect of a decrease in toxic concentration on respiratory outcomes. To do so, I use a wide and rich dataset on property prices. I find first a positive effect of an improvement in air pollution on housing prices. While the purchase of cheap flats does not reflect this positive and significant effect, people buying more expensive flats are willing to pay more for an improvement in air quality. Hedonic approach may not always reflect the environmental health risks.

The first results indicate that SO2, even at levels below current air quality standards in most of the world, has significant negative impacts on the severity of a respiratory disease, suggesting that the strengthening of regulations on SO2 pollution would yield additional benefits. The second part of the project suggests buyers of cheap flat are less sensitive to pollution than other due to budget constraint.

General conclusions

There is widespread concern that the increasing concentration of pollutants in the environment may have detrimental health impacts. In this context, the main objective of this thesis has been to shed light on social inequalities linked to the health consequences of air pollution.

We propose to investigate the causal effect of atmospheric pollution on mortality rate, infants' health and respiratory outcomes in France using fixed effect strategy, natural experiments and a unique dataset combining data on health, weather, socioeconomic status, environmental quality and property prices.

Epidemiological field studies most often fail to control for factors that may confound the effect of exposure to pollutants. It is for example highly likely that poorer households and households that care less for health outcomes may migrate to more polluted areas, a phenomenon called Tiébout sorting in economics (Banzhaf and Walsh, 2008). Low income and weak concern for health have direct effects on health and may thus confound the effect of pollution and bias the estimates of the health consequences of environmental pollution. This thesis is part of a growing literature in applied economics aiming at using "‘natural experiments"’ to control for possible confounding factors and identify causal effects.

We combine variations in the exposure to pollution in the air both across space and across time. Across space, we know that individuals living in different places are exposed to different levels of pollution at a given point in time. Individuals living close to the industrial or transport areas where air pollution is important are more at risk of breathing air pollutants than individuals living further away. It remains possible, though, that Tiébout sorting worse health outcomes in areas more exposed to air pollution, even when pollution is not sprayed. In order to gauge the effects of these potential confounding factors, we propose to use differences in health outcomes between these two groups when air pollution is decreasing. This difference in difference strategy yields credible estimates of the causal effects of air pollution on health under testable assumptions.

This thesis emphazises that relatively low concentrations of air pollutants, at levels below the regulated threshold, are related to a range of adverse health effects. SO2 in particular has significant negative impacts on the severity of a respiratory disease and infant health. At risks population are particularly impacted by SO2: children below 5 years old and people over 70 years old. In this context, strengthening regulation about air pollution should be part of policy debates. In addition,

results are strongly supportive of the hypothesis that NO2 has a positive impact on mortality rate with the effect being larger when considering the subsample of departments with the highest level of pollution. Women seems also to be relatively more impacted by NO2.
In terms of socioeconomic results, we acknowledge that health disparities exist among French departments but seem to be more related to socioeconomic factors than differences in sensitivity to pollution. People located near refineries may be more exposed to pollution and might be more adversely health affected to pollution policies may be effective at the margin.Besides, budget constraint plays a crucial role in the willingness to pay for an improvement in air quality. Even if people are aware and sensitive to an improvement in air quality, their budget constraint may prevent them to respond to an air quality improvement. In terms of health, the hedonic approach may not always reflect the true environmental health risks.

This findings are consistent with the results of international studies that have examined the relationship between economic inequality, environmental quality and health. It also confirms the importance of ambient air pollution and reinforces the need for politicians to take into account environmental justice in France. Policy-makers regulating industries and transport have to weigh-in potential benefits from reduced health costs from air pollution. Localized pollution policies may be especially effective at the margin. Further studies of air pollution and socioeconomic status in Europe are useful in order to better inform this important public policy process. Future research should focus for instance on longer term effect of air pollution concentration on health outcomes.

Bibliography

Agency, Environmental Protection. 2011. *Sulfur Dioxide @ONLINE*.

Agency, Environmental Protection. 2012a. *Control Emmissions Technologies - Transport Dispersion of Air Pollutants @ONLINE*.

Agency, Environmental Protection. 2012b. *Sulfur Dioxide @ONLINE*.

Almond, Douglas, & Currie, Janet. 2011. Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, **25**(3), 153–72.

Almond, Douglas, Edlund, Lena, & Palme, Marten. 2009. Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. *The Quarterly Journal of Economics*, **124**(4), 1729–1772.

Bajari, Patrick, Fruehwirth, Jane Cooley, il Kim, Kyoo, & Timmins, Christopher. 2012. A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution. *American Economic Review*, **102**(5), 1898–1926.

Beatty, Timothy K.M., & Shimshack, Jay P. 2011. School buses, diesel emissions, and respiratory health. *Journal of Health Economics*, **30**(5), 987 – 999.

Besson, Danielle. 2006. *Consommation de tabac : la baisse s est accentuée depuis 2003*. Tech. rept.

Bickel, P., Friedric, R., Droste-Frank, B., Bachman, T. M., & Gressman, Rabl, A. 2005. *ExternE, Externalities of Energy, Methodology 2005 update*. Tech. rept.

Bishai, D., & Opuni, M. 2009. Are infant mortality rate declines exponential? The general pattern of 20th century infant mortality rate decline. *Popul Health Metr*, **7**.

Blanchard, O. Summers, L. 1986. *Hysteresis And The European Unemployment Problem*. Tech. rept.

Bolte, G., Tamburlini, G., & Kohlhuber, M. 2010. Environmental inequalities among children in Europe-evaluation of scientific evidence and policy implications. *Eur J Public Health*, **20**(1), 14–20.

Brookshire, David S, et al. 1982. Valuing Public Goods: A Comparison of Survey and Hedonic Approaches. *American Economic Review*, **72**(1), 165–77.

Cambois, Emmanuelle, & Jusot, Florence. 2010. Contribution of lifelong adverse experiences to social health inequalities: findings from a population survey in France. *The European Journal of Public Health*.

Campagna, Dave, Lefranc, AgnÃĺs, Nunes-Odasso, Catherine, & Ruth, Ferry. 2003.

Evaluation des risques de la pollution urbaine sur la santé en Ile-de-France (erpurs) : liens avec la mortalité 1987-1998.

Chanel, O., Scapecchi, P., & Vergnaud, J.-C. 2000. Pollution atmospherique et sante : evaluation monetaire et effets de long terme.

Chay, Kenneth, Dobkin, Carlos, & Greenstone, Michael. 2003. The Clean Air Act of 1970 and Adult Mortality. *Journal of Risk and Uncertainty*, **27**(3), 279–300.

Chay, Kenneth Y., & Greenstone, Michael. 2003. The Impact Of Air Pollution On Infant Mortality: Evidence From Geographic Variation In Pollution Shocks Induced By A Recession. *The Quarterly Journal of Economics*, **118**(3), 1121–1167.

Chay, Kenneth Y., & Greenstone, Michael. 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, **113**(2), 376–424.

Clougherty, J. E. 2010. A Growing Role for Gender Analysis in Air Pollution Epidemiology. *Environmental Health Perspectives*, **118**(2), 167–176.

Crocker, T.D., Horst, R.L., & Gillette, D.G. 1977. *Oxidant Air Pollution and Work Performance of Citrus Harvest Labor*. US Socioconomic environmental series. U.S. Environmental Protection Agency, Office of Research and Development, Health Effects Research Laboratory.

Cropper, Maureen L. 2000. Has Economic Research Answered the Needs of Environmental Policy? *Journal of Environmental Economics and Management*, **39**(3), 328–350.

Currie, Janet, & Neidell, Matthew. 2004. *Air Pollution and Infant Health: What Can We Learn From California's Recent Experience.* NBER Working Papers 10251. National Bureau of Economic Research, Inc.

Currie, Janet, & Walker, Reed. 2011. Traffic Congestion and Infant Health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, **3**(1), 65–90.

Currie, Janet, Hanushek, Eric A, Kahn, E. Megan, Neidell, Matthew, & Rivkin, Steven G. 2009. Does Pollution Increase School Absences? *The Review of Economics and Statistics*, **91**(4), 682–694.

Currie, Janet, Greenstone, Michael, & Moretti, Enrico. 2011. Superfund Cleanups and Infant Health. *American Economic Review*, **101**(3), 435–41.

Currie, Janet, Davis, Lucas, Greenstone, Michael, & Walker, Reed. 2013. Do Housing Prices Reflect Environmental Health Risks? Evidence from More than 1600 Toxic Plant Openings and Closings.

Davis, Lucas W. 2004. The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster. *American Economic Review*, **94**(5), 1693–1704.

Deguen, S., & Zmirou, D. 2012. Social inequalities resulting from health risks related to ambient air quality; A European review. *European Journal of Public Health*, **20**(1), 27–35.

Delucchi, Mark, Murphy, James, & McCubbin, Donald. 2002 (Feb.). *The Health and Visibility Cost of Air Pollution: A Comparison of Estimation Methods*. Institute of Transportation Studies, Working Paper Series qt03s2x9xb. Institute of Transportation Studies, UC Davis.

Deschenes, Olivier, Greenstone, Michael, & Guryan, Jonathan. 2009. Climate Change and Birth Weight. *American Economic Review*, **99**(2), 211–17.

Ettner, Susan L. 1996. New evidence on the relationship between income and health. *Journal of Health Economics*, **15**(1), 67–85.

Finkelstein, Murray M, Jerrett, Michael, DeLuca, Patrick, Finkelstein, Norm, Verma, Dave K, Chapman, Kenneth, & Sears, Malcolm R. 2003. Relation between income, air pollution and mortality: a cohort study. *CMAJ*, **169**(5), 397–402.

Fiscella, K., & Franks, P. 2000. Individual Income, Income Inequality, Health, and Mortality: What Are the Relationships? **35**, 307âĂŞ18.

Frankenberg, Elizabeth, McKee, Douglas, & Thomas, Duncan. 2005. Health Consequences of Forest Fires in Indonesia. *Demography*, **42**(1), pp. 109–129.

Graff Zivin, Joshua, & Neidell, Matthew. 2013. Environment, Health, and Human Capital. *Mimeo, Columbia University.*

Greenstone, Michael, & Gallagher, Justin. 2008. Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. *The Quarterly Journal of Economics*, **123**(3), 951–1003.

Greenstone, Michael, & Hanna, Rema. 2011. *Environmental Regulations, Air and Water Pollution, and Infant Mortality in India.* Working Papers 1114. Massachusetts Institute of Technology, Center for Energy and Environmental Policy Research.

Grineski, Sara, Bolin, Bob, & Boone, Christopher. 2007. Criteria Air Pollution and Marginalized Populations: Environmental Inequality in Metropolitan Phoenix, Arizona. *Social Science Quarterly*, **88**(2), 535–554.

Groot, Wim, & Maassen van den Brink, Henriette. 2007. The health effects of education. *Economics of Education Review*, **26**(2), 186–200.

Grossman, M. 2000. The human capital model. in Handbook of Health Economics. **1A**.

Gunasekara, Fiona Imlach, Carter, Kristie, & Blakely, Tony. 2011. Change in income and change in self-rated health: Systematic review of studies using repeated measures to control for confounding bias. *Social Science & Medicine*, **72**(2), 193–201.

Hanna, Brid Gleeson. 2007. House values, incomes, and industrial pollution. *Journal of Environmental Economics and Management*, **54**(1), 100–112.

Hanna, Rema, & Oliva, Paulina. 2011. *The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City*. Working Paper 17302. National Bureau of Economic Research.

Harper, Krista, Steger, Tamara, & Filcak, Richard. 2009. Environmental justice and Roma communities in Central and Eastern Europe. *Environmental Policy and Governance*, **19**(4), 251–268.

Hokby, Stina, & Soderqvist, Tore. 2003. Elasticities of Demand and Willingness to Pay for Environmental Services in Sweden. *Environmental & Resource Economics*, **26**(3), 361–383.

INSEE. 2010. *L'inscription territoriale de la raffinerie des Flandres dans la région Nord-Pas-de-Calais*. Tech. rept.

Janke, Katharina, Propper, Carol, & Henderson, John. 2009. Do current levels of air pollution kill? The impact of air pollution on population mortality in England. *Health Economics*, **18**(9), 1031–1055.

Joyce, Theodore, Grossman Michael, & Goldman, Fred. 1988a. An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health. *Journal of Urban Economics*, **25**(2), 32–51.

Joyce, Theodore, Grossman Michael, & Goldman, Fred. 1988b. An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health. *Journal of Urban Economics*, **25**(2), 32–51.

Katsouyanni, K, Zmirou, D, Spix, C, Sunyer, J, Schouten, J P, Ponka, A, Anderson, H R, Le Moullec, Y, Wojtyniak, B, & Vigotti, M A.

Krupnick, Alan, Alberini, Anna, Cropper, M, Simon, N, O'Brien, B, Goeree, R, & Herntzelman, M. 2002. Age, Health and the Willingness to Pay for Mortality Risk

Reductions: A Contingent Valuation Survey of Ontario Residents. *Journal of Risk and Uncertainty*, **24**(2), 161–86.

Kuminoff, Nicolai V., Parmeter, Christopher F., & Pope, Jaren C. 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of Environmental Economics and Management*, **60**(3), 145–160.

Kunzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, F., V., Puybonnieux-Texier Jr, & Quenel, P. 2000. Public-Health Impact of Outdoor and Traffic-related Air Pollution: A European Assessment. *Lancet*, **356**, 795–801.

Laigle, L., & Oehler, V. Les enjeux sociaux et environnementaux du developpement urbain : la question des inegalites ecologiques.

Lancaster, KJ. 1966. A new approach to consumer theory. *Journal of Political Economy*, **74**, 132– 157.

Laurent, eloi. 2011. Issues in environmental justice within the European Union. *Ecological Economics*, **70**(11), 1846–1853.

Laurent, O., Bard, D., Filleul, L., & Segala, C. 2007. Effect of socioeconomic status on the relationship between atmospheric pollution and mortality. *J Epidemiol Community Health*, **61**(8), 665–75.

Lindahl, M. 2005. Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income. **40**, 144âĂŞ68.

Maantay, J. 2007. Asthma and air pollution in the Bronx: methodological and data considerations in using GIS for environmental justice and health research. *Health Place*, **13**(1), 32–56.

Maesano, A., Moreau, D., Caillaud, D., Lavaud, F., Le Moullec, Y., Taytard, A., Pauli, G., & Charpin, D. 2007. Residential Proximity Fine Particles Related to Allergic Sensitisation and Asthma in Primary School Children. *Respiratory Medicine*, **101**(8), 1721–1729.

Marr, Linsey C. Matthew E. R. 2010. Effect of Air Pollution on Marathon Running Performance. *Medicine Science in Sports Exercise*, **42**(3), 585.

Martens, W.J. 1998. Climate change, thermal stress and mortality changes. *Social Science amp; Medicine*, **46**(3), 331 – 344.

Maslianskaia Pautrel, Masha. 2009. La valorisation de la qualite de l'air par l'approche hedonique : une revue de la litterature. *Revue franÃğaise d'economie*, **23**(3), 109–160.

McLeod, H., Langford, I. H., Jones, A. P., Stedman, J. R., Day, R. J., Lorenzoni, I., & Bateman, I. J. 2000. The relationship between socio-economic indicators and air pollution in England and Wales: implications for environmental justice. *Regional Environmental Change*, **1**, 78–85.

Midorikawa, K., Uchida, T., Okamoto, Y., Toda, C., Sakai, Y., Ueda, K., Hiraku, Y., Murata, M., Kawanishi, S., & Kojima, N. 2004. Metabolic activation of carcinogenic ethylbenzene leads to oxidative DNA damage. *Chem Biol Interact*, **150**(3), 271–81.

Mohai, P., M., Paula, Lantz, J., Morenoff, J., & P., Richard. 2009. Racial and socioeconomic disparities in residential proximity to polluting industrial facilities: evidence from the Americans' Changing Lives Study. *American Journal of Public Health*, **99**(3), S649–56.

Morello Frosch, R., Pastor, M J.R, & Sadd, J. 2002. Integrating Environmental Justice and the Precautionary Principle in Research and Policy Making: The Case of Ambient Air Toxics Exposures and Health Risks among Schoolchildren in Los Angeles. *Annals of the American Academy of political and social science*, **584**, 47–68.

Moretti, Enrico, & Neidell, Matthew. 2011. Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles. *Journal of Human Resources*, **46**(1), 154–175.

Neidell, Matthew. 2009. Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations. *Journal of Human Resources*, **44**(2).

Neidell, Matthew J. 2004. Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, **23**(6), 1209–1236.

OMS. 2006. Lignes directrices OMS relatives Ãă la qualite de l'air : particules, ozone, dioxyde d'azote et dioxyde de soufre.

OMS. 2009. La pollution de l'air Ãă l'interieur des habitations.

OMS. 2012. Environmental health inequalities in Europe.

Organization, World Health. 2005. *Effects of air pollution on childrens̀ health and development.* Report.

Parry, Ian, & Small, Kenneth. 2002 (Mar.). *Does Britain or the United States Have the Right Gasoline Tax?* Discussion Papers dp-02-12-. Resources For the Future.

Parry, Ian W. H., & Small, Kenneth A. 2005. Does Britain or the United States Have the Right Gasoline Tax? *American Economic Review*, **95**(4), 1276–1289.

Pascal, L., Blanchard, M, Fabre, P.I., Larrieu, S., Borreli, D.and Host, S., Chardon, B., Chatignoux, E., Prouvost, H., Jusot, J.F., Wagner, V., Declercq, C., Medina, S., & Lefranc, A. 2009. *Liens a court terme entre la mortalité et les admissions a l´ hôpital et les niveaux de pollution atmosphérique dans neuf villes francaises.* Tech. rept.

Pastor, Manuel, Sadd, Jim, & Hipp, John. 2001. Which Came First? Toxic Facilities, Minority Move-In, and Environmental Justice. *Journal of Urban Affairs*, **23**(1), 1–21.

Pearce, Jamie R, Richardson, Elizabeth A, Mitchell, Richard J, & Shortt, Niamh K. 2010. Environmental justice and health: the implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom. *Transactions of the Institute of British Geographers*, **35**(4), 522–539.

Peng, Roger, Welty, Leah, & McDermott, Aidan. 2004 (July). *The National Morbidity, Mortality, and Air Pollution Study Database in R.* Johns Hopkins University Dept. of Biostatistics Working Paper Series 1044. Berkeley Electronic Press.

Pierre Lecoq, Frédéric Dor, Cécile Kairo. 2009. *Description des valeurs repères toxicologiques utilisées lors dèxpositions aiguës par inhalation des populations*. Report. The French Institute for Public Health Surveillance, InVS.

Portney, Paul R. 1981. Housing prices, health effects, and valuing reductions in risk of death. *Journal of Environmental Economics and Management*, **8**(1), 72 – 78.

Premji, S., Bertrand, F., Smargiassi, A., & Daniels, M. 2007. Socioeconomic Correlates of Municipal Level Pollution Emissions on Montreal Island. *Canadian Journal of Public Health*, **98**(2), 138–42.

Rabl, A. 2003. *Mortalité due a la pollution de l air: comment interpréter les résultats.*

Rady, Sven. 1998. *Boom In, Bust Out: Young Households and the Housing Price Cycle.* FMG Discussion Papers dp310. Financial Markets Group.

Ransom, Michael R, & Iii, C. Arden Pope. 1995. External Health Costs Of A Steel Mill. *Contemporary Economic Policy*, **13**(2), 86–97.

Rosen, Sherwin. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, **82**(1), 34–55.

Sanders, Nicholas J., & Stoecker, Charles F. 2011 (Sept.). *Where Have All the Young Men Gone? Using Gender Ratios to Measure Fetal Death Rates*. NBER Working Papers 17434. National Bureau of Economic Research, Inc.

Schikowski, T., Sugiri, D., Reimann, V., Pesch, B., Ranft, U., & Ursula, K. 2008. Air Pollution Exposure in Urban Areas to Social Differences in Respiratory Health. *BMC public health*, 8–179.

Soleille, Sébastien. 2004. *Lìndustrie du raffinage et le devenir des fiouls lourds*. Report. French national institute for industrial environment and risks, INERIS.

Stringhini, Silvia, Berkman, Lisa, Dugravot, Aline, Ferrie, Jane E., Marmot, Michael, Kivimaki, Mika, & Singh-Manoux, Archana. 2012. Socioeconomic Status, Structural and Functional Measures of Social Support, and Mortality: The British Whitehall II Cohort Study, 1985âĂŞ2009. *American Journal of Epidemiology*.

Subramanian, S.V., & Kawachi, I. 2000. Being well and doing well: on the importance of income for health. **15**, 13–22.

Touloumi, G., Katsouyanni, K., Zmirou, D., Schwartz, J., Spix, C., Ponce de Leon, A., Tobias, A., Quennel, P., Rabczenko, D., Bacharova, L., Bisanti, L., Vonk, J. M., & Ponka, A. 1997. Short-term Effects of Ambient Oxidant Exposure on Mortality: A Combined Analysis within the APHEA Project. *American Journal of Epidemiology*, **146**(2), 177–185.

Viel, Jean-FranÃğois, Hagi, Mathieu, Upegui, Erika, & Laurian, Lucie. 2011. Environmental justice in a French industrial region: Are polluting industrial facilities equally distributed? *Health / Place*, **17**(1), 257 – 262. Health Geographies of Voluntarism.

Wilhelm, Michelle, Qian, Lei, & Ritz, Beate. 2009. Outdoor air pollution, family and neighborhood environment, and asthma in {LA} {FANS} children. *Health Place*, **15**(1), 25 – 36.

World Health Organization (WHO). 2011. *Air quality and health @ONLINE.*

Yamada, E., Vanna, A.T., Naspitz, C.K., & Sole, D. 2002. International Study of Asthma and Allergies in Childhood (ISAAC): validation of the written questionnaire (eczema component) and prevalence of atopic eczema among Brazilian children. *J Investig Allergol Clin Immunol*, **12**(1), 34–41.

Zabel, Jeffrey E., & Kiel, Katherine A. 2000. Estimating the Demand for Air Quality in Four U.S. Cities. *Land Economics*, **76**(2), 174–194.

Zivin, Joshua Graff, & Neidell, Matthew. 2013 (Apr.). *Environment, Health, and Human Capital.* NBER Working Papers 18935. National Bureau of Economic Research, Inc.